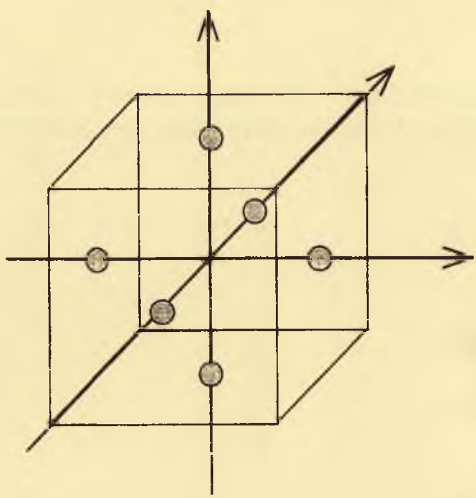
Міністерство освіти і науки України
Вінницький державний технічний університет

В.В. МОТИГІН, С.М. ПАВЛОВ

ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ В ІНЖЕНЕРНИХ ДОСЛІДЖЕННЯХ

(лабораторний практикум)



Вінниця ВДТУ 2001

Міністерство освіти і науки України
Вінницький державний технічний університет

В.В. Мотигін, С.М.Павлов

ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ В ІНЖЕНЕРНИХ ДОСЛІДЖЕННЯХ

(лабораторний практикум)

Затверджено Ученою радою Вінницького державного технічного університету як навчальний посібник для студентів технічних спеціальностей. Протокол №9 від 27 квітня 2000 р.

Вінниця ВДТУ 2001

УДК 519.6
М 85

Р е ц е н з е н т и:
Р.Н. Квєтний, доктор технічних наук, професор
В.О. Поджаренко, доктор технічних наук, професор
А.А. Козак, кандидат технічних наук

Рекомендовано до видання Ученою радою Вінницького державного технічного університету Міністерства освіти і науки України

Мотигін В.В., Павлов С.М.
М 85 **Планування експерименту в інженерних дослідженнях** (лабораторний практикум). Навчальний посібник. - Вінниця: ВДТУ, 2001.- 82 с.

В навчальному посібнику розглянуто деякі питання математичної теорії оптимального експерименту та методика використання її в експериментальних дослідженнях, також приведено завдання та порядок виконання лабораторних робіт. Зокрема розглянуті основні характеристики і експериментальний аналіз випадкових величин; повний та дробовий факторні експерименти; метод випадкового балансу; дисперсійний аналіз.
Навчальний посібник призначено для студентів технічних спеціальностей.

УДК 519.6

Зміст

ВСТУП

Експеримент є найважливішою частиною наукового дослідження. Впровадження сучасних експериментально-статистичних методів дозволяє звести до мінімуму інтуїтивний підхід до організації експерименту. Він замінюється науково обгрунтованою програмою проведення експериментального дослідження.

Основна мета більшості експериментальних досліджень полягає в знаходженні такої сукупності вхідних керованих змінних, при яких цільова функція, що оптимізується, приймає екстремальне значення. Ця мета досягається за допомогою мінімально можливого числа дослідів при мінімумі затрат часу і коштів.

Навіть при неповному знанні внутрішніх закономірностей явищ, які вивчаються в об'єктах дослідження, шляхом направленого експерименту можна одержати математичну модель складного об'єкта, що включає найбільш суттєві фактори. Така математична модель може використовуватися не тільки для управління і знаходження необхідних режимів функціонування, але і для виявлення взаємозв'язків в об'єкті, що раніше були невідомі.

Навчальний посібник призначений для теоретичного вивчення та практичного закріплення основних розділів математичної теорії планування експерименту, що вивчається в дисциплінах "Теорія планування експерименту" та "Основи науково-дослідної роботи" для студентів технічних спеціальностей.

Для проведення лабораторних занять необхідний персональний комп'ютер типу IBM з операційною системою MS DOS 6.22.

Прикладна програма Black Box(BB) моделює складні об'єкти дослідження, що описуються статистичними рівняннями у вигляді поліномів не вище другого порядку з декількома вхідними та вихідними величинами і схильних до впливу випадкових завад.

Програма BB використовується для вивчення різноманітних експериментально-статистичних методів отримання математичного опису об'єкта дослідження та його подальшої оптимізації.

Програма BB включає файли:

BB.BAT – командний файл для запуску програми із завантаженням резидентного калькулятора.;

BB.EXE – власне програма;

BB.HLP – файл допомоги;

MOD1.TAB, MOD2.TAB, MOD3.TAB, MOD4.TAB – довідкові таблиці. (Зміст усіх файлів, окрім BB.EXE, може змінюватись користувачем.)

Робота з програмою здійснюється у такий спосіб. Після запуску програми необхідно ввести номер потрібної моделі об'єкта (від 1 до 12), а потім – значення середньоквадратичного відхилення (СКВ) (від 1 до 99) у відповідні знакомісця (рис. 1.). При цьому можна використовувати клавіші:

0,...,9 - для набору значення;

left, right – "клавіші-стрілки" для переміщення на 1 позицію ліворуч або праворуч;

Home,End або PgUp, PgDn - для переміщення до початку та до кінця вікна вводу;

Enter - для підтвердження (закінчення) вводу. Здійснюється перехід до наступного вікна вводу;

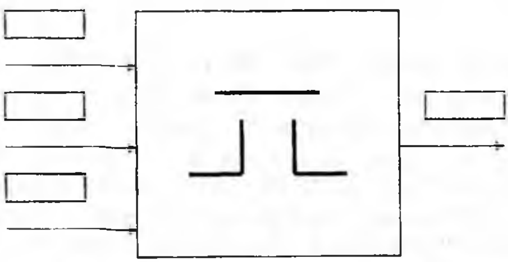
Esc - для відмови від вводу (зі збереженням попереднього значення);

Insert - для перемикання режимів: вставка / заміна;

Del - для видалення символу в поточній позиції;

Backspace - для видалення символу ліворуч від курсора;

Ввести значення вхідних сигналів і натиснути F9



F2 - Таблиця рівномірно розподілених випадкових чисел (0..99)
F3 - Значення критерію Кохрена (q=0,05)
F4 - Значення критерію Стьюдента
F5 - Значення критерію Фішера (q=0,05)

F9-Go! F1-HELP F2..F5-Довідкові дані Alt+C-Calc F8-Model F10-Exit

Рисунок 1 – Вигляд екрана при запуску програми BlackBox

Down (стрілка вниз), Tab - для підтвердження (закінчення) вводу та переходу до наступного вікна вводу (діє аналогічно Enter).

Примітка. Можливе введення значень тільки з діапазону, який вказаний на екрані в дужках.

Після вводу потрібних значень з'являється наступний екран із зображенням «чорного ящика», який відповідає рис 1.

Необхідно ввести або змінити значення сигналів на входах «ящика» (при потребі), для чого, додатково до вже описаних, можна використовувати клавішу PgUp (стрілка вгору) – для повернення (переходу) до попереднього входу моделі.

Для отримання результату на виході слід натиснути клавішу **F9**, після чого на екрані буде натиснута «кнопка» (супроводжується звуковим сигналом) та біля зображення виходу (виходів) з'явиться відповідне (відповідні) значення. Далі можна знов натиснути F9, зберігаючи на входах попередні значення, або змінити їх. Після закінчення роботи з обраною моделлю можна або зовсім вийти з програми (клавіша F10), або повернутись до вибору моделі (клавіша F8).

В ході роботи на екран видаються підказки про необхідні дії, а в його нижньому рядку при цьому відображаються доступні функціональні клавіші.

З метою полегшення обробки результатів експерименту в програмі передбачений виклик резидентного калькулятора (ALT+C), а також чотирьох довідкових таблиць:

- таблиці рівномірно розподілених випадкових чисел (0..99);
- таблиці значень критерію Кохрена (q=0,05);
- таблиці значень критерію Стьюдента;
- таблиці значень критерію Фішера (q=0,05);

Ці довідкові дані знаходяться у файлах MOD1.TAB, MOD2.TAB, MOD3.TAB та MOD4.TAB. Для їх виведення на екран монітору при роботі програми слугують клавіші F2...F5. При натисканні F1 на екран виводиться вікно допомоги (файл BB.HLP), де описане призначення клавіш у програмі. Для перегляду Help або довідкових таблиць використовуються стрілки вверх та вниз. Повернення до звичайного режиму роботи програми здійснюється натисканням Enter або Esc.

Тема 1 Основні характеристики і експериментальний аналіз випадкових величин

Мета – *вивчення основних понять теорії імовірності, ознайомлення з основними характеристиками випадкових величин і можливих засобів їхнього експериментального визначення.*

КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

1.1 Елементи теорії імовірності

Випадковою величиною називають таку величину, значення якої змінюються при повторенні дослідів деяким, заздалегідь не передбаченим способом. На відміну від невипадкових, детермінованих величин для випадкової величини не можна заздалегідь точно сказати, яке конкретно значення вона прийме в визначених умовах, а можна тільки зазначити закон її розподілу. Закон розподілу вважається заданим, якщо:

- зазначена множина можливих значень випадкової величини;
- зазначений спосіб кількісного визначення можливості попадання випадкової величини в будь-яку область множини можливих значень.

Імовірність потрапляння в задану область може бути визначена таким чином

$$P_m = \lim_{N \to \infty} \frac{N_m}{N}. \tag{1.1}$$

де N_m – кількість спостережень випадкової величини, що опинилась в даній області; N – загальна кількість спостережень (частотне визначення імовірності).

Аналітичні вирази законів розподілу $F(x)$ випадкової величини X показують імовірність того, що випадкова величина не перевищує деякого заданого або поточного значення x, тобто $F(x) = P\{X \le x\}$. З цього випливає, що імовірність того, що значення випадкової величини X розташоване між x_1 і x_2, рівна різниці значень функції розподілу, обчислених в цих двох точках

$$P\{x_1 < X \le x_2\} = F(x_2) - F(x_1). \tag{1.2}$$

Аналогічно,

$$P\{X > x\} = 1 - F(x). \tag{1.3}$$

Інтегральній функції розподілу випадкової величини X притаманні такі властивості:

1) $\lim\limits_{x \to \infty} F(x) = F(-\infty) = 0;$

2) $\lim\limits_{x \to \infty} F(x) = F(\infty) = 1;$

3) $F(x) \geq 0$ для всіх x;

4) $F(x_2) \geq F(x_1)$, якщо $x_2 > x_1$.

Далі будемо розглядати лише безперервні випадкові величини, область можливих значень яких включає всі точки з деякого інтервалу. Для такої величини можливий вигляд розподілу $F(x)$ зображений на рис. 2.

Якщо функцію $F(x)$ можливо диференціювати для всіх значень випадкової величини X, то закон розподілу імовірностей може бути виражений в аналітичній формі також за допомогою диференціальної функції розподілу імовірностей:

$$f(x) = \frac{dF(x)}{dx} = \lim_{\Delta x \to \infty} \frac{P\{x < X \leq x + \Delta x\}}{\Delta x} (\Delta x > 0). \qquad (1.4)$$

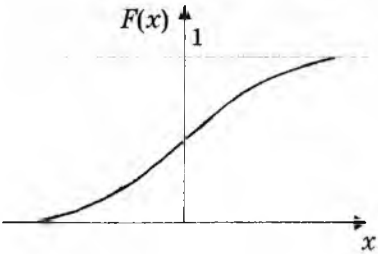


Рисунок 2 – Інтегральна функція розподілу випадкової величини

Таким чином значення функції $f(x)$ приблизно рівне відношенню імовірності потрапляння випадкової величини в інтервал $(x,\ x + \Delta x)$ до довжини Δx цього інтервалу, коли Δx – нескінченно мала величина. Тому функцію $f(x)$ називають також функцією щільності розподілу імовірностей (або коротше – функцією щільності імовірності).

Відмітимо основні властивості функції $f(x)$:

1) $f(x) \geq 0$;

2) $\int\limits_{-\infty}^{\infty} f(x)dx = 1$;

3) $\int\limits_{-\infty}^{x} f(z)dz = F(x)$;

4) $\lim\limits_{x \to \infty} f(x) = 0$ (z – змінна інтегрування).

За допомогою диференціальної функції розподілу обчислюється імовірність розміщення випадкової величини в будь-якій області з множини її можливих значень. Зокрема:

$$P\{X \leq a_1\} = \int\limits_{-\infty}^{a_1} f(x)dx, P\{X > a_2\}, P\{X > a_2\} = \int\limits_{a_2}^{\infty} f(x)dx,$$

$$P\{a_1 < X \leq a_2\} = \int\limits_{a_1}^{a_2} f(x)dx. \qquad (1.5)$$

Для безперервної випадкової величини імовірність можна визначити як відносну частину площі під кривою щільності розподілу імовірностей $f(x)$. Так, наприклад, імовірність того, що випадкова величина X приймає значення, менші a_1, дорівнює відносній частині площі під кривою $f(x)$ зліва від точки a_1 (рис. 3,а); імовірність того, що ця величина X прийме значення, більше a_2, рівна відносній частині площі під кривою $f(x)$ справа від точки a_2 (рис. 3,б); імовірність того, що вона прийме значення, що знаходиться між a_1 і a_2, рівна відносній частині площі під кривою $f(x)$ між точками a_1 і a_2 (рис. 3,в).

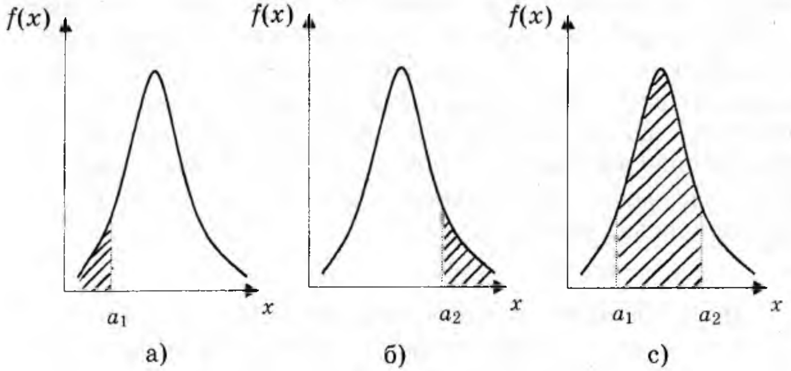


Рисунок 3 – Диференціальна функція розподілу імовірності випадкової величини

Як інтегральна так і диференціальна функції розподілу є вичерпними імовірними характеристиками випадкової величини. Але деякі основні властивості випадкових величин можуть бути описані простіше за допомогою визначених числових параметрів. Найбільшу роль серед них на практиці відіграють два параметри, що характеризують центр розсіювання (центр розподілу) випадкової величини і ступінь її розсіювання навколо цього центру. Найбільш розповсюдженою характеристикою центру розподілу є математичне очікування m_x випадкової величини X (часто називають також генеральним середнім значенням)

$$m_x = \int_{-\infty}^{\infty} x f(x) dx. \tag{1.6}$$

Ступінь розсіювання випадкової величини X відносно m_x може бути охарактеризована за допомогою генеральної дисперсії $\sigma_x{}^2$

$$\sigma_x^2 = \int_{-\infty}^{\infty}(x - m_x)^2 f(x)dx. \quad (1.7)$$

Якщо $f(x)$ все більше концентрується поблизу m_x, то значення σ_x^2 зменшується. Якщо ж мають місце досить відмінні від m_x значення випадкової величини X і для них $f(x)$ не дуже мала, то дисперсія σ_x^2 збільшується. Квадратний корінь з дисперсії σ_x^2 називають середнім квадратичним відхиленням σ_x.

Часто для описання практичної ситуації є необхідним використання одночасно декількох (в простих випадках – двох) випадкових величин. Для завдання ймовірнісних властивостей двох випадкових величин X, Y використовують двовимірні (сумісні) функції розподілу імовірностей: інтегральна $F(x, y)$ і диференціальна $f(x, y)$. Функція $F(x, y)$, що характеризує імовірність того, що перша початкова величина приймає деяке значення, менше або рівне x, а друга – значення, що менше або рівне y, називають інтегральною функцією сумісного розподілу двох випадкових величин

$$F(x, y) = P\{ X \le x;\ Y \le y\}. \quad (1.8)$$

Як і для однієї безперервної випадкової величини, якщо функція $F(x, y)$ достатньо гладка, то її можна продиференціювати, в результаті отримаємо двовимірну диференціальну функцію розподілу імовірностей (двовимірна щільність імовірності)

$$f(x, y) = \frac{\partial^2}{\partial x \partial y} F(x, y). \quad (1.9)$$

Функція $f(x, y)$ має такі властивості:

1) $f(x, y) \ge 0$;

2) $\int_{-\infty}^{\infty}\int f(x, y)dxdy = 1$;

3) $F(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(z_1, z_2)dz_1dz_2$;

4) $\lim\limits_{\substack{|x| \to \infty \\ |y| \to \infty}} f(x, y) = 0$.

(z_1, z_2 – змінні інтегрування).

Імовірність того, що випадкові величини X, Y одночасно потраплять в деяку умовну область Ω, складає

$$P\{x, y \in \Omega\} = \iint_{\Omega} f(x, y)dxdy. \quad (1.10)$$

Також,

$$P\{a_1 < X \le a_2; b_1 < Y \le b_2\} = \int_{a_1}^{a_2}\int_{b_1}^{b_2} f(x,y)dxdy. \qquad (1.11)$$

По відомій двовимірній щільності $f(x, y)$ легко знайти одновимірні функції розподілу $f(x)$, $f(y)$ кожної випадкової величини

$$f(x) = \int_{-\infty}^{\infty} f(x,y)dy, f(y) = \int_{-\infty}^{\infty} f(x,y)dx. \qquad (1.12)$$

Дві випадкові величини X, Y називають незалежними, якщо

$$f(x,y) = f(x)f(y). \qquad (1.13)$$

Як і в одновимірному випадку, основні властивості двовимірної сукупності величин X, Y можуть бути охарактеризовані за допомогою ряду чисельних параметрів. При цьому в якості параметрів, що описують поведінку кожної з випадкових величин окремо, як і в розглянутих вище, використовуються математичне очікування і дисперсія відповідної випадкової величини: m_x, m_y, σ_x^2, σ_y^2. Крім подібного роду параметрів для двовимірної сукупності можуть бути побудовані параметри, що характеризують ступінь взаємозалежності змінних X і Y. Найпростішими з них є коваріація двох випадкових величин (яку називають також кореляційним моментом)

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{\infty}\int f(x - m_x)(y - m_y)f(x,y)dxdy, \qquad (1.14)$$

а також нормований показник зв'язку – коефіцієнт кореляції

$$\rho_{xy} = \frac{\mathrm{cov}(X,Y)}{\sigma_x \sigma_y}. \qquad (1.15)$$

За своїм фізичним змістом коефіцієнт кореляції є далеко не вичерпною характеристикою статистичного зв'язку, характеризуючи лише ступінь лінійної залежності між X та Y. Коефіцієнт кореляції змінюється в межах $-1 \le \rho_{xy} \le 1$. Якщо $\rho_{xy}=1$, то випадкові величини мають додатню кореляцію, тобто $y=a_0+a_1x$, де a_0 і a_1 – постійні, причому $a_1>0$. Якщо ж $\rho_{xy}=-1$, то випадкові величини мають від'ємну кореляцію, тобто $y=a_0-a_1x$. Якщо $\rho_{xy}=0$, то випадкові величини X та Y не корельовано: $a_1=0$. В тому випадку, коли X та Y – незалежні випадкові величини, для них $\rho_{xy}=0$; тобто, вони не корельовані. Зворотне твердження в загальному випадку невірне: X та Y можуть бути пов'язані навіть функціонально і все ж мати нульовий коефіцієнт кореляції (при цьому, дійсно, функціональний зв'язок повинен бути нелінійним).

Всі описані вище функції і пов'язані з ними параметри є теоретичними, що характеризують визначені властивості об'єкта, що вивчається. На практиці майже завжди ці характеристики невідомі і виникає задача експериментального (емпіричного) визначення тих чи інших характеристик випадкових величин на основі спостережень.

1.2 *Експериментальний аналіз одновимірної сукупності*

Нехай є набір (вибірка) експериментальних даних x_1, x_2, ..., x_N. Опрацювання цих даних для одержання емпіричних характеристик одновимірної випадкової величини роблять звичайно в такій послідовності:

1. Побудова варіаційного ряду. Варіаційний ряд z_1, z_2, ..., z_N одержують із вихідних даних шляхом розташування x_m (m = 1, 2, ..., N) у порядку зростання від x_{min} до x_{max}.

2. Побудова діаграми накопичених частот $\hat{F}_N(x)$, що є емпіричним аналогом інтегрального закону розподілу. Діаграму будують у відповідності з формулою

$$\hat{F}_N(x) = \sum_{i=1}^{\mu_N(x)} \frac{1}{N}, \tag{1.16}$$

де $\mu_N(x)$ – кількість елементів в виборці, для яких значення $x_j<x$. Практично це роблять таким чином. На осі абсцис вказують значення спостережень x_m (або z_l). Значення по осі ординат дорівнює нулю лівіше точки x_{min}; в точці x_{min} і далі в усіх інших точках x_m діаграма має стрибок, рівний $1/N$. Якщо є декілька значень x_m, що збігаються, то в цьому місці на діаграмі виникає стрибок, рівний λ/N, де λ – кількість точок, що збігаються. Зрозуміло, що для величин $x>x_{max}$ значення діаграми накопичених частот рівне 1. Відмітимо, що якщо $N\to\infty$, то $\hat{F}_N(x) \to F(x)$.

Використовуючи дані попереднього прикладу, побудуємо відповідну діаграму (рис. 4).

3. Побудова гістограми вибірки. Гістограма $\hat{f}_N(x)$ є емпіричним аналогом функції щільності розподілу $f(x)$. Звичайно її будують таким чином:

- Знаходять попередню кількість квантів (інтервалів), на яку повинна бути розбита вісь Ox. Цю кількість K визначають за допомогою формули
$$K=1+3{,}2\lg N, \tag{1.17}$$

де знайдене значення округляють до найближчого цілого числа.

- Визначають довжину інтервалу

$$\Delta x=(x_{max}-x_{min})/K. \qquad (1.18)$$

розмір Δx можливо дещо округлити для зручності обчислень.

- Середину області зміни вибірки (центр розподілу) $(x_{max}+x_{min})/2$ приймають за центр деякого інтервалу, після чого легко знаходять межі й остаточну кількість зазначених інтервалів так, щоб у сукупності вони перекривали всю область від x_{min} до x_{max}.

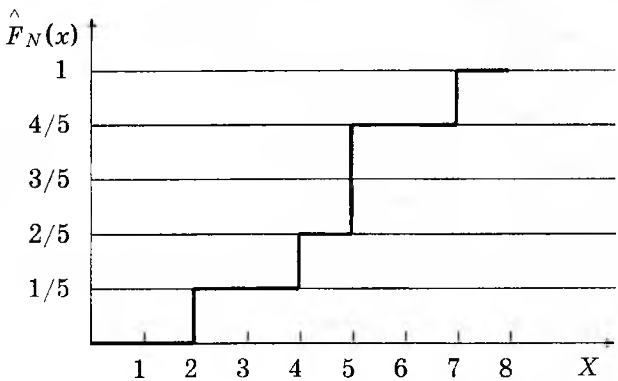


Рисунок 4 – Діаграма накопичених частот

- Підраховують кількість спостережень N_m, що потрапили в кожний квант: N_m дорівнює числу членів варіаційного ряду, для яких справедлива нерівність

$$x_m \le z_1 < x_m + \Delta x. \qquad (1.19)$$

Тут x_m і $x_m+\Delta x$ – межа m-го інтервалу. Значення z_1, що потрапило на межу між $(m-1)$-м і m-м інтервалами, відносять до m-го інтервалу.

- Підраховують відносну кількість (відносну частоту) спостережень N_m/N, що потрапили в даний квант.
- Будують гістограму, що представляє собою ступінчату криву, значення на якій на m-му інтервалі $(x_m,\ x_m+\Delta x)(m=1,2,\ ...,\ K)$ постійне і дорівнює N_m/N, або з урахуванням умови $\int_{-\infty}^{+\infty} f(z)dz=1$ дорівнює $N_m/N)\Delta x$.

Нехай, наприклад, є вибірка з 40 спостережень, а варіаційний ряд, що їй відповідає має вигляд $x_{min}=z_1=0,3$; $z_2=0,4;...;$ $z_{10}=x_{max}=7,1$.

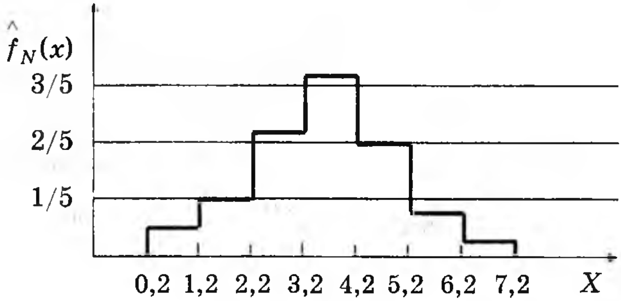


Рисунок 5 – Гістограма вибірки

За формулою (1.17) отримуємо K = 1+3,2*lg40 = 1+3.2*1.602 = 6,13; приймаємо K=7. Тоді Δx = (7,1–0,3)/7 = 0,971; вибираємо Δx =1. Знаходимо $(x_{min} + x_{max})/2$ = (0,3+7,1)/2 = 3,7, після чого легко визначаємо межі інтервалів (рис. 5). Припустимо, що після такого розбиття визначили, що в перший інтервал потрапило два значення x_i: $N_1=2$; в другий – чотири: $N_2=4$; в наступні: $N_3=9$; $N_4=13$; $N_5=8$; $N_6=3$; $N_7=1$. Імовірна гістограма зображена на рис. 5.

4. Визначення оцінок математичного очікування $\bar{x}$, дисперсії S_x^2 і середнього квадратичного відхилення S_x роблять за формулами:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i; \qquad (1.20)$$

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left(x_i - \bar{x}\right)^2; \qquad (1.21)$$

$$s_x = +\sqrt{s^2}. \qquad (1.22)$$

Відмітимо, що множник $1/(N-1)$ (замість $1/N$) в формулі (1.21) вводять для отримання незміщеної оцінки дисперсії.

1.3 Експериментальний аналіз двовимірної сукупності

Нехай отримана вибірка з двовимірної сукупності при спостереженні двох випадкових величин X і Y (табл. 1.1).

Таблиця 1.1

i	1	2	3	...	j	...	N
X	x_1	x_2	x_3	...	x_i	...	x_N
Y	y_1	y_2	y_3	...	y_i	...	y_N

Опрацювання результатів спостережень у даному випадку можна здійснювати за такою схемою:

1. Побудова поля розсіювання, як правило, перший крок при опрацюванні результатів спостережень двовимірної сукупності випадкових величин X,Y; для цього на площині з координатами x, y відмічають експериментальні точки.

2. Складання таблиці двовимірного розподілу. Цю таблицю складають таким чином. Осі Ox і Oy розбивають на окремі інтервали довжиною Δx і Δy. Величини Δx, Δy, їх кількість K_x, K_y і розміщення цих інтервалів для кожної зі змінних X і Y знаходять за допомогою відомих правил. Відповідні межі наносять на діаграму розсіювання (рис. 6) і потім підраховують кількість точок $N_{m1,m2}$, що потрапили в кожний з прямокутників, що виникли (за умови, що якщо будь-яка точка розташована на межі, то її відносять до правого або верхнього прямокутника). Далі складають таблицю (табл. 1.2). У цій таблиці відзначають величини $N_{m1,m2}$, а також відносні розміри $N_{m1,m2}/N$ (нижче діагоналі). Подібну таблицю можна використовувати як вихідну для побудови гістограм і діаграм накопичених частот у тривимірному просторі, що є емпіричними аналогами двовимірного інтегрального закону і двовимірної функції розподілу.

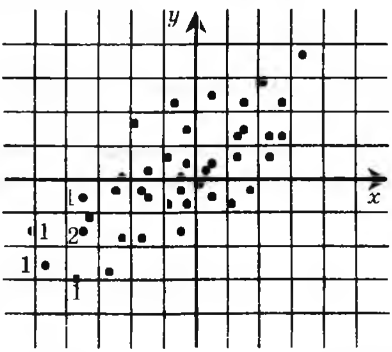


Рисунок 6 – Поле розсіювання

За допомогою таблиці двовимірного розподілу легко одержати вихідні дані для побудови гістограм, що відповідають кожній з одновимірних випадкових величин X і Y. Для цього достатньо скласти значення таблиці або кожного стовпця (при побудові гістограми для Y), або по кожному рядку (при побудові гістограми для X).

3. Обчислення оцінки коефіцієнта кореляції роблять за формулою

$$\hat{\rho}_{xy} \frac{1}{(N-1)s_x s_y} \sum_{i=1}^{N} (x_i - \bar{x})(y_i - \bar{y}), \qquad (1.23)$$

де $\bar{x}$, $\bar{y}$ і s_x, s_y знаходять за допомогою формул (1.20) – (1.22). Відмітимо, що між значенням і знаком коефіцієнта кореляції, з одного боку, і видом діаграми розсіювання – з іншого, існує певний зв'язок. Якщо початок координат перенести в точку $(\bar{x};\bar{y})$, тоді: 1) при $\rho > 0$ точки на діаграмі розсіювання групуються в основному в I і III квадрантах, а при $\rho < 0$ – у II і IV; 2) при $\rho \approx 0$ точки безладно розкидані у всіх чотирьох квадрантах; при $\rho = \pm 1$ точки групуються на прямих (які знаходяться або в I і III квадрантах; або II і IV).

Таблиця 1.2

Інтервали $y \rightarrow$	$(y_1, y_1+\Delta y)$	$(y_2, y_2+\Delta y)$	…	$(y_{m1}, y_{m2}+\Delta y)$	…	$(y_{k2}, y_{k2}+\Delta y)$
Інтервали $x \downarrow$			…		…	
$(x_1, x_1+\Delta x)$	0					1
		0	…		…	1/40
$(x_2, x_2+\Delta x)$	0					
		0	…		…	
…	…	…		…		…
$(x_{m1}, x_{m2}+\Delta x)$			…	N_{m1}, m_2	…	
				$N_{m1}, m_2/N$		
…	…	…	…	…	…	…
$(x_{k1}, x_{k2}+\Delta x)$	1	1	…		…	
	1/40	1/40				

ЗАВДАННЯ І ПОРЯДОК ВИКОНАННЯ РОБОТИ

1. Досліджувати властивості одновимірної випадкової величини x_1 (модель об‘єкту 1):

- зняти вибірку, що складається з 30 значень випадкової величини x_1, при заданому середньоквадратичному відхиленні σ_x=15 для значення x_{10}=55;
- побудувати варіаційний ряд;
- за допомогою варіаційного ряду побудувати діаграму накопичених частот;
- побудувати гістограму вибірки;
- розрахувати оцінки математичного очікування, дисперсії і середньоквадратичного відхилення випадкової величини x_1.

2. Досліджувати властивості двовимірної сукупності випадкових величин x_1, y_1(модель об'єкту 2):

- встановити σ_y= 20, x_{10}=8,5;
- виконати той же експеримент при σ_y=5;
- виконати той же експеримент при σ_y=1;
- за отриманими даними побудувати три поля розсіювання;
- скласти таблиці двовимірних розподілів;
- побудувати гістограму для випадкової величини y_1 (за даними або п.2а, або п.2б, або п.2в – за вказівкою викладача);
- обчислити для усіх трьох випадків оцінки коефіцієнтів кореляції.

3. Розрахунок коефіцієнтів кореляції робити за незгрупованими даними. Переконатися, що між виглядом поля розсіювання, величиною і знаком коефіцієнта кореляції є певний зв'язок.

ФОРМУЛИ ДЛЯ РОЗРАХУНКУ

1. $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$. 2. $S_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2$. 3. $S_x = +\sqrt{S^2}$.

4. $P = \frac{1}{(N-1)S_x S_y}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Дати визначення випадкової величини.
2. Дати визначення одновимірного інтегрального і диференціального законів розподілу випадкової величини і назвати їх властивості.
3. Дати визначення двовимірного інтегрального і диференціального законів розподілу випадкових величин і назвати їх властивості.
4. Які числові параметри найбільш часто використовуються в якості мір розташування і розсіювання одновимірної та двовимірної сукупностей випадкових величин?
5. Яким чином проводиться побудова варіаційного ряду, діаграми накопичених частот, гістограми вибірки одновимірної випадкової величини?
6. Яким чином проводиться побудова поля розсіювання й упорядкування таблиці розподілу двовимірної сукупності випадкових величин?

Тема 2 ПОВНИЙ ТА ДРОБОВИЙ ФАКТОРНІ ЕКСПЕРИМЕНТИ

Мета – *вивчення методів планування експериментів для отримання лінійної та неповної степеневої математичних моделей статики складних об'єктів.*

ТЕОРЕТИЧНІ ВІДОМОСТІ

2.1 Елементи регресивного аналізу

Статистичні методи планування активного експерименту є одним із емпіричних способів отримання математичного опису статики складних об'єктів дослідження, тобто рівняння зв'язку відгуку об'єкта та незалежних керованих нормованих вхідних змінних(факторів) $\vec{z}^{\mathrm{T}} = (z_1, z_2, ..., z_n)$. При цьому математичний опис подається у вигляді деякого поліному – відрізка ряду Тейлора, в який розкладається невідома залежність в межах основної точки $\vec{z}_0$

$$M\{y\} = \varphi(z_1, z_2, \ldots, z_n) = \beta_0 + \sum_{i=1}^{n} \beta_i z_i + \sum_{\substack{i;l=1 \\ i<l}}^{n} \beta_{il} z_i z_l + \sum_{i=1}^{n} \beta_{ii} z_i^2 + \ldots, \quad (2.1)$$

де $\beta_i = \dfrac{\partial\varphi}{\partial z_i}\Big|_{\vec{z}=\vec{z}_0}$; $\beta_{il} = \dfrac{\partial^2\varphi}{\partial z_i^2}\Big|_{\vec{z}=\vec{z}_0}$; $\beta_{ii} = \dfrac{1}{2}\dfrac{\partial^2\varphi}{\partial z_i^2}\Big|_{\vec{z}=\vec{z}_0}$ – теоретичні коефіцієнти.

Внаслідок присутності некерованих та навіть неконтрольованих факторів зміна величини y носить випадковий характер, тому функціональна залежність $\varphi(\vec{z})$ не дає точного зв'язку між керованими факторами та математичним очікуванням випадкової величини y

$$M\{y_g\} = \varphi(\vec{z}_g). \quad (2.2)$$

Тут $\vec{z}_g^{\mathrm{T}} = (z_{1g}, z_{2g}, \ldots, z_{ng})$ – g-та точка простору незалежних керованих факторів (факторного простору). В такому випадку за результатами експерименту можливо відшукати рівняння регресії в формі деякого поліному

$$\hat{y} = b_0 + \sum_{i=1}^{n} b_i z_i + \sum_{\substack{i;l=1 \\ i<1}}^{n} b_{il} z_i z_l + \sum_{i=1}^{n} b_{ii} z_i^2 + \ldots, \quad (2.3)$$

де вибіркові коефіцієнти регресії є лише оцінками для теоретичних коефіцієнтів, відповідно $\beta_0, \beta_i, \beta_{ii}, \ldots$, а $\hat{y}$ – оцінкою для $M\{y\}$. Нехай $\vec{z}_g (g = 1,2,\ldots,N)$ – точки факторного простору, в яких проводиться експеримент. Тоді задача пошуку оцінок коефіцієнтів рівняння регресії (2.3) за результатами досліду в N точках факторного простору є типовою задачею множинного регресивного аналізу в тому випадку, якщо виконуються рекомендації:

1. Результати спостережень $y_1, y_2, \ldots, y_N$ відгуку в N точках факторного простору, представляють собою незалежні нормально розподілені випадкові величини, тобто на них діють нормально розподілені випадкові завади ε з нульовим математичним очікуванням $M\{\varepsilon\} = 0$.

2. Дисперсії $\sigma^2\{y_g\}\,(g = 1,2,\ldots,N)$ рівні. Це означає, що отримані при проведенні багаторазових повторних спостережень над величиною y_g в точках $\bar{z}_g$ вибіркові оцінки $s_g^2\{y\}$ однорідні, дисперсія ж $\sigma^2\{y_g\}$ не залежить від математичного очікування $M\{y_g\}$, тобто не відрізняється від дисперсії $\sigma^2\{y_g\}$, отриманої при повторних спостереженнях в будь-якій іншій точці $\bar{z}_g$ факторного простору (відтворення з рівною точністю).

3. Незалежні керовані фактори z_1, z_2, ..., z_n вимірюються з несуттєво малими помилками в порівнянні з помилкою в визначенні y (мається на увазі вплив цих помилок на величину y в порівнянні з впливом некерованих та неконтрольованих факторів ε).

Розглянемо задачу знаходження коефіцієнтів рівняння регресії (2.3) на прикладі рівняння другого порядку з чотирма незалежними факторами, при цьому, звичайно, вся процедура та зроблені висновки можуть бути поширені на рівняння будь-якого степеня з n незалежними факторами. Насамперед спростимо систему позначень: введемо фіктивну зміну $z_0=1$, та позначимо $z_0=f_0$; $z_1=f_1$; ...; $z_4=f_4$; $z_{12}=f_5$; ...; $z_4{}^2=f_8$; $z_1z_2=f_9$; ...; $z_3z_4=f_{14}$. В новій системі позначень поліном другого степеня записується як лінійне однорідне рівняння

$$\hat{y} = \sum_{j=0}^{14} b_j f_j. \qquad (2.4)$$

Нехай проводяться досліди в N точках 4-факторного простору χ, причому y_g – значення відгуку при досліді в точці $\bar{z}_g\,(g = 1,2,\ldots,N)$. Коефіцієнти рівняння (2.4) знаходяться на основі методу найменших квадратів, тобто, з умови мінімуму суми квадратів відхилень значень відгуку $\hat{y}_g$, передбачених рівнянням (2.4) для умов дослідів в точках $\bar{z}_g$, від значень y_g, що спостерігаються при дослідах в цих точках,

$$\sum_{g=1}^{N}\left(y_g - \hat{y}_g\right)^2 = \sum_{g=1}^{N}\left[y_g - \sum_{j=0}^{14} b_j f_{gj}\right]^2. \qquad (2.5)$$

Оскільки задача полягає у пошуку значень коефіцієнтів b_j, що мінімізують вираз (2.5), вона вирішується за допомогою системи так званих нормальних рівнянь, отриманих порівнянням до нуля частинних похідних від квадратичної форми (2.5) по змінних параметрах $b_j (j=0, 1, 2, ..., 14)$. Система звичайних рівнянь має вигляд

$$\left.\begin{array}{l} c_{0,0}b_0 + c_{0,1}b_1 + \ldots + c_{0,14}b_{14} = \alpha_0 \\ c_{1,0}b_0 + c_{1,1}b_1 + \ldots + c_{1,14}b_{14} = \alpha_1 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ c_{14,0}b_0 + c_{14,1}b_1 + \ldots + c_{14,14}b_{14} = \alpha_{14} \end{array}\right\}, \qquad (2.6)$$

$$\mathbf{C} = \left\{c_{jl}\right\} = \begin{pmatrix} c_{0,0} & c_{0,1} \cdots c_{0,14} \\ c_{1,0} & c_{1,1} \cdots c_{11,14} \\ \ldots & \ldots\ldots\ldots \\ c_{14,0} & c_{14,1} \cdots c_{14,14} \end{pmatrix} \qquad (2.7)$$

С – матриця коефіцієнтів системи (2.6), елементи якої знаходяться таким чином

$$c_{jl} = \sum_{g=1}^{N} f_{gj} f_{gl}. \qquad (2.8)$$

Матриця $\mathbf{F} = \left\{f_{gj}\right\}$ називається матрицею незалежних змінних, тоді $\mathbf{F}^{\mathrm{T}} = \left\{f_{jg}\right\}$ – матриця, отримана транспонуванням матриці **F**, причому

$$\mathbf{F} = \left\{f_{gj}\right\} = \begin{pmatrix} f_{1,0} & f_{1,1} \cdots f_{1,14} \\ f_{2,0} & f_{2,1} \cdots f_{2,14} \\ \ldots & \ldots\ldots\ldots \\ f_{N,0} & f_{N,1} \cdots f_{N,14} \end{pmatrix}. \qquad (2.9)$$

Можна показати, що $\mathbf{C}=\mathbf{F}^{\mathrm{T}}\mathbf{F}$.

Вільні члени α_j системи нормальних рівнянь визначаються за допомогою рівності

$$\alpha_{jl} = \sum_{g=1}^{N} f_{gj} y_g. \qquad (2.10)$$

Для того, щоб система (2.6) мала єдиний розв'язок, необхідно і достатньо, щоб матриця **С** була невиродженою, тобто її визначник повинен бути відмінним від нуля $|\mathbf{C}| \neq 0$. Неважко показати, що ця умова зводиться до умови лінійної незалежності вектор-стовпців матриці **F**, тобто, для того щоб система (2.6) мала єдиний розв'язок, необхідно та достатньо, щоб вектор-стовпці матриці **F** були лінійно незалежні. Аналіз розв'язку системи (2.6), отриманого, наприклад, за формулою Крамера

$$b_l = |\mathbf{C}_l|/|\mathbf{C}| \quad (l = 0,1,2,\ldots,14), \tag{2.11}$$

де $|\mathbf{C}_l|$ – визначник, який отримується з $|\mathbf{C}|$ при заміні елементів c_{jl} l-го стовпця відповідними вільними членами α_j показує, що значення коефіцієнтів b_l залежать від кількості членів рівняння регресії, тобто збільшення або зменшення числа членів рівняння впливає на значення коефіцієнтів всіх включених в рівняння членів. Така невизначеність в оцінці коефіцієнтів регресії вкрай ускладнює їх фізичну інтерпретацію. В тому випадку, коли матриця C діагональна, тобто виконується умова

$$c_{jl} = \sum_{g=1}^{N} f_{gj} f_{gl} = 0 \text{ при } j \neq l, \tag{2.12}$$

система (2.6) розпадається на незалежні нормальні рівняння

$$\left.\begin{array}{l} c_{0,0}b_0 = \alpha_0 \\ c_{1,1}b_1 = \alpha_1 \\ \ldots\ldots\ldots\ldots \\ c_{14,14}b_{14} = \alpha_{14} \end{array}\right\}, \tag{2.13}$$

розв'язок яких знаходиться з співвідношень

$$b_j = \alpha_j / c_{jj} \quad (j = 0,1,2,\ldots,14). \tag{2.14}$$

При цьому вдається позбутися від невизначеності, пов'язаної з неоднозначним оцінюванням коефіцієнтів регресії. Таким чином, для отримання незалежних один від одного оцінок коефіцієнтів регресії потрібно спланувати експеримент так, щоб виконувалися умови лінійної незалежності та ортогональності вектор-стовпців матриці **F** незалежних змінних, або, як будемо її називати, матриці планування (МП).

2.2 Повний факторний експеримент

Для побудови лінійних та неповних степеневих математичних моделей застосовують повний факторний експеримент, який має ортогональну матрицю планування. Математичне описання поверхні відгуку об'єкта навколо точки базового режиму $\bar{x}_0^{\mathrm{T}} = (x_{10}; x_{20}; \ldots; x_{n0})$ можливо отримати варіюванням кожного з факторів x_i на двох рівнях, які відрізняються від базового рівня x_{i0} на величину інтервалу варіювання Δx_i. Інтервал варіювання по кожному керованому фактору вибирають так, щоб приріст величини відгуку y до базового значення y_0 при реалізації $x_{i0}\pm\Delta x_i$ можна було виділити на фоні "шуму" при невеликому числі паралельних дослідів.

Повним факторним експериментом (ПФЕ) називається експеримент, який реалізує всі можливі неповторні комбінації рівнів n незалежних керованих факторів, кожний з яких варіює на двох рівнях. Число цих комбінацій $N=2^n$ визначає тип ПФЕ. Для спрощення подальше викладення побудуємо на прикладі планування типу $N=2^3$, тобто на прикладі об'єкта з трьома (n=3) незалежними керованими факторами x_1, x_2, x_3. При плануванні експерименту проводять перетворення розмірних керованих незалежних факторів x_i в безрозмірні, нормовані

$$z_i = \left(x_i - x_{i0}\right)/\Delta x_i \; ; \qquad (2.15)$$

це дає можливість легко побудувати ортогональну МП та значно полегшує подальші розрахунки, оскільки в цьому випадку верхні та нижні рівні варіювання $z_{iв}$ та $z_{iн}$ у відносних одиницях дорівнюють відповідно +1 та −1 незалежно від фізичної природи факторів, значень основних рівнів та інтервалів варіювання факторів Δx_i.

Якщо для трифакторної задачі теоретичне рівняння регресії відносно нормованих факторів має вигляд

$$M\{y\} = \beta_0 + \sum_{i=1}^{3} \beta_i z_i + \sum_{\substack{i;l=1 \\ i<l}}^{3} \beta_{il} z_i z_l + \beta_{123} z_1 z_2 z_3 \qquad (2.16)$$

(тобто степенями факторів вище першого можна знехтувати), тоді ПФЕ дає можливість знайти роздільні оцінки коефіцієнтів β_i. Оскільки зміна вихідної величини y носить випадковий характер, то є можливість визначити лише вибіркові коефіцієнти регресії b_i, b_{il} для оцінювання теоретичних коефіцієнтів β_i, β_{il}. Процес знаходження моделі методом ПФЕ складається з: планування експерименту; проведення експерименту на об'єкті дослідження; перевірки відтворення (однорідності вибіркових дисперсій s_g^2) експерименту; отримання математичної моделі об'єкта з перевіркою статистичної значимості вибіркових коефіцієнтів регресії; перевірки адекватності математичного описання.

Планування експерименту. Матрицю планування ПФЕ можна представити у вигляді табл. 2.1. Її складають за такими правилами:

1. Кожен g-тий рядок матриці містить набір координат z_{ig} точки, в якій проводиться g-й дослід (i=1, 2, ..., n; g=1, 2, ..., N).

2. Як вказувалося вище, вводять фіктивну змінну z_0=+1.

3. Оскільки змінні z_i приймають лише значення +1 та −1, всі взаємодії $z_i z_l$ (i; l=1, 2, 3; $i \neq l$) можуть приймати тільки такі ж значення.

4. В першому рядку (g=1) всі керовані фактори вибирають на нижньому рівні, тобто $z_i=-1$. Наступні g-ті варіанти варіювання при складанні МП вибирають так: при порядковому переборі всіх варіантів частота зміни знаку факторів для кожного наступного фактора z_{i+1} вдвічі менша, ніж для попереднього z_i (див. табл. 2.1). Три стовпця керованих факторів утворюють план експерименту (обведено жирною лінією), а інші стовпці МП отримуються перемноженням відповідних значень керованих факторів і необхідні для розрахунку відповідних коефіцієнтів при взаємодіях.

План ПФЕ типу 2^4 (n=4) можна побудувати або вказаним вище способом, або на базі плану ПФЕ типу 2^3, повторивши його двічі: один раз – при величині $z_4=-1$, другий раз – при $z_4=+1$. Аналогічно можуть бути отримані плани для як завгодно великого числа n незалежних керованих факторів.

Аналізуючи табл. 2.1 помітимо, що матриця планування має лінійно незалежні, попарно ортогональні вектор-стовпці, звідки випливає діагональність матриці **C** коефіцієнтів системи нормальних рівнянь (2.6), тобто, взаємна незалежність оцінок коефіцієнтів рівнянь регресії, не говорячи вже про простоту розрахунку цих оцінок. Необхідно відмітити, що в тих випадках, коли вплив на відгук змінних типу z_i^2 стає суттєвим і з'являється необхідність обчислення оцінок коефіцієнтів при них, ПФЕ не дає можливості визначити роздільні оцінки таких коефіцієнтів, як b_0, b_i (i=1, 2, ..., n), оскільки відповідні вектор-стовпці МП рівні між собою, а тому ця матриця має лінійно залежні вектор-стовпці.

Таблиця 2.1 – Матриця планування ПФЕ 2^3

g	z_0	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
1	+1	−1	−1	−1	+1	+1	+1	−1
2	+1	+1	−1	−1	−1	−1	+1	+1
3	+1	−1	+1	−1	−1	+1	−1	+1
4	+1	+1	+1	−1	+1	−1	−1	−1
5	+1	−1	−1	+1	+1	−1	−1	+1
6	+1	+1	−1	+1	−1	+1	−1	−1
7	+1	−1	+1	+1	−1	−1	+1	−1
8	+1	+1	+1	+1	+1	+1	+1	+1

Проведення експерименту на об'єкті дослідження. Оскільки зміна відгуку y носить випадковий характер, то в кожній точці $\bar{x}_g$ приходиться проводити m паралельних дослідів та результати спостережень усереднювати

$$\bar{y}_g = \frac{1}{m}\sum_{k=1}^{m} y_{gk}. \tag{2.17}$$

Нехай у випадку який розглядається, m=3. Перед реалізацією плану на об'єкті необхідно рандомізувати варіанти варіювання факторів, тобто, за допомогою таблиці рівномірно розподілених випадкових чисел визначити послідовність реалізації варіантів варіювання плану в $N\times m$ дослідах. Рандомізацію проводять таким чином. В таблиці рівномірно розподілених випадкових чисел (див. додаток А) вибирають будь-який стовпець з якого по черзі беруть числа від 1 до $N\times m$ і записують послідовно в m стовпців k_1, k_2, ..., k_m (кожне число береться тільки один раз). Якщо в одному стовпці таблиці додатку А не виявилось усіх $N\times m$ потрібних чисел, переходять до наступного її стовпця. Нехай, наприклад, k_1=8 при g=3, це значить, що третій варіант варіювання реалізується в експерименті восьмим за порядком. Результати спостереження за експериментом відповідно до варіантів варіювання плану записують в стовпці $y_{gk_1}, y_{gk_2}, y_{gk_3}$.

Перевірка відтворюваності експерименту є не що інше як перевірка виконання другої передумови регресивного аналізу про однорідність вибіркових дисперсій s_g^2. Завдання полягає в перевірці, під час дослідів, гіпотези про рівність генеральних дисперсій $\sigma^2\{y_1\} = \sigma^2\{y_2\} = \ldots = \sigma^2\{y_N\}$ відповідно в точках $\bar{x}_1$, $\bar{x}_2$, ..., $\bar{x}_N$. Оцінки дисперсій знаходять за формулою

$$s_g^2 = \frac{1}{m-1}\sum_{k=1}^{m}\left(y_{gk} - \bar{y}_g\right)^2. \tag{2.18}$$

Оскільки всі оцінки дисперсій отримані за вибірками однакового об'єкта m=3, то кількість ступенів свободи для всіх однакова і складає

$$\nu_{1\,ВІДТ} = m - 1. \tag{2.19}$$

В цьому випадку для перевірки гіпотези про однорідність оцінок s_g^2 дисперсій потрібно користуватись критерієм Кохрена, який оснований на законі розподілу відношення максимальної оцінки дисперсії до суми всіх зіставлених оцінок дисперсій, тобто

$$G = \frac{\max\limits_{g}\left\{s_g^2\right\}}{\sum\limits_{g=1}^{N} s_g^2\{y\}}. \tag{2.20}$$

Якщо обчислене за даними експерименту (емпіричне) значення критерію G виявиться менше за критичне значення $G_{кр}$, знайденого за таблицею (див. додаток Б) для $\nu_{1ВІДТ}=m-1$ і $\nu_{2ВІДТ}=N$ (в даному випадку $\nu_{1ВІДТ}=2$ і $\nu_{2ВІДТ}=8$) та вибраного рівня значимості $q_{ВІДТ}[\%]$ (звичайно 5%), то гіпотеза про однорідність вибіркових дисперсій відповідає результатам спостережень. При цьому всю групу вибіркових дисперсій s_g^2 можна вважати оцінками для однієї і тієї ж генеральної дисперсії $\sigma^2\{y\}$ відтворюваності експерименту, звідки найкраща її оцінка має вигляд

$$s^2_{ВІДТ}\{y\}=\frac{1}{N}\sum_{g=1}^{N}s_g^2\{y\} \tag{2.21}$$

з кількістю ступенів свободи

$$\nu_{зн}=N(m-1). \tag{2.22}$$

Якщо перевірка відтворюваності досліду дала негативний результат, тоді залишається визнати його невідтворюваність відносно керуючих факторів внаслідок присутності несприятливих флуктуацій некерованих та неконтрольованих факторів. При цьому необхідно або збільшити кількість паралельних дослідів для варіантів варіювання з більшими значеннями вибіркових дисперсій s_g^2, або використовувати в подальшому модифікацію методу найменших квадратів, придатну при невиконанні передумови про відтворюваність експерименту.

Отримання математичної моделі об'єкта. Як вже вказувалось раніше, користуючись методом ПФЕ, можна отримати опис об'єкта, що вивчається, в формі (2.6). При ПФЕ отримуються незалежні оцінки b_0, b_i, b_{il} відповідних коефіцієнтів $\beta_0, \beta_i, \beta_{il}$, тобто $b_0\to\beta_0$, $b_1\to\beta_1$, $b_2\to\beta_2$, $b_3\to\beta_3$. Ці оцінки легко знайти за формулами:

$$b_0=\frac{1}{N}\sum_{g=1}^{N}z_{0g}\bar{y}_g,\qquad b_i=\frac{1}{N}\sum_{g=1}^{N}z_{ig}\bar{y}_g\quad (i=1,2,\ldots,n),$$

$$b_{il}=\frac{1}{N}\sum_{g=1}^{N}z_{ig}z_{lg}\bar{y}_g\quad (i;\,l=1,2,\ldots,n;\ i\neq l). \tag{2.23}$$

Після визначення оцінок b коефіцієнтів регресії необхідно перевірити гіпотези про їх значущість, тобто перевірити відповідні нуль-гіпотези $\beta=0$. Перевірку таких гіпотез проводять за допомогою критерію Ст'юдента, емпіричне значення якого

$$t=|b|/s\{b\}, \tag{2.24}$$

$$s^2\{b\}=\frac{1}{Nm}s^2_{відт}\{y\} \tag{2.25}$$

де $s^2\{b\}$ – дисперсія оцінки коефіцієнта b; N – кількість точок факторного простору, в яких проводиться експеримент; m – кількість паралельних дослідів в цих точках. Якщо знайдена величина t перевищує значення $t_{кр}$, що знайдене з додатку В для кількості ступенів $\nu_{зн} = N(m-1)$, при заданому рівні значущості $q_{зн}$ (звичайно 5%), тобто $\operatorname{sign}(t-t_{кр}) = +1$, то перевірну нуль-гіпотезу H_0:$\beta=0$ відкидають і відповідну оцінку коефіцієнта визнають значущою. В іншому випадку, тобто при $\operatorname{sign}(t-t_{кр}) = -1$, нуль-гіпотезу не відкидають і оцінку b_i вважають статично незначущою, тобто $\beta = 0$.

Статистична незначимість оцінки коефіцієнта регресії b_i може бути обумовлена такими причинами:

1) даний i-й фактор не має функціонального зв'язку з відгуком y, тобто $\beta_i=0$;

2) рівень x_{i0} базового режиму $\bar{x}_0$ знаходиться в точці окремого екстремуму функції відгуку по фактору x_i і тоді $\beta_i = \frac{\partial y}{\partial z_i} = 0$;

3) інтервал варіювання Δx_i вибраний маленьким;

4) внаслідок впливу некерованих і неконтрольованих факторів велика помилка відтворюваності експерименту.

Ортогональне планування дозволяє визначати довірчі границі незалежно для кожного коефіцієнта регресії. Тому, якщо якась з оцінок коефіцієнтів виявиться незначною, то її можна відкинути без перераховування всіх інших. Після цього математичну модель об'єкта складають у вигляді рівняння зв'язку відгуку y і факторів z_i, які включають тільки значимі оцінки коефіцієнтів.

Перевірка адекватності математичного опису. Щоб перевірити гіпотезу про адекватність математичного опису дослідним даним, достатньо оцінити відхилення передбачуваної за отриманим рівнянням регресії величини відгуку $\hat{y}_g$ від результатів спостереження $\bar{y}_g$ в одних і тих же g-х точках факторного простору. Розсіювання результатів спостережень поблизу рівняння регресії, що оцінює істинну функцію відгуку, можна характеризувати за допомогою дисперсії адекватності

$$s^2_{АД} = \frac{m}{N-d}\sum_{g=1}^{N}\left(\bar{y}_g - \hat{y}_g\right)^2, \qquad (2.26)$$

де d - кількість членів апроксимуючого поліному. Дисперсія адекватності визначається за кількістю ступенів свободи

$$\nu_{АД} = N - d. \qquad (2.27)$$

Перевірка гіпотези про адекватність складається в виявленні співвідношення між дисперсією адекватності $s^2_{АД}$ та оцінкою дисперсії відтворюванності відгуку $s^2_{ВІДТ}$. Якщо ці оцінки дисперсії однорідні, то математичний опис адекватно представляє результати дослідів, якщо ж ні, то опис вважається неадекватним. Перевірку гіпотези про адекватність проводять з використанням F-критерію Фішера. Критерій Фішера дозволяє перевірити гіпотезу про однорідність двох вибіркових дисперсій $s^2_{АД}$ і $s^2_{ВІДТ}\{y\}$. В тому випадку, якщо $s^2_{АД} > s^2_{ВІДТ}\{y\}$, F-критерій характеризується відношенням

$$F = s^2_{А} / s^2_{ВІДТ}\{y\}. \qquad (2.28)$$

Якщо обраховане за результатами спостережень емпіричне значення критерію F менше критичного $F_{кр}$, знайденого із додатка Г для відповідних ступенів свободи:

$$\nu_{1АД} = N - d, \quad \nu_{2АД} = \nu_{ЗН} = N(m - 1) \qquad (2.29)$$

при заданому рівні значущості $q_{АД}$, то гіпотезу про адекватність не відкидають. В іншому випадку гіпотезу відкидають і математичний опис визнається неадекватним.

Перевірка адекватності можлива при $\nu_{1АД}>0$. Якщо кількість N варіантів варіювання плану ПФЕ дорівнює кількості всіх значущих оцінок коефіцієнтів регресії ($N=d$), то для перевірки гіпотези про адекватність математичного опису, ступенів свободи не залишається ($\nu_{1АД}=0$). Якщо ж деякі оцінки коефіцієнтів регресії виявились незначущими, то число d членів рівняння регресії в цьому випадку менше числа N варіантів варіювання ($N>d$) та для перевірки гіпотези про адекватність залишається одна чи декілька ступенів свободи ($\nu_{1АД}>0$).

В тому випадку, коли гіпотеза про адекватність відкидається необхідно перейти до більш складної форми математичного опису або, проводити дослід з меншим інтервалом варіювання. Слід відмітити, що максимальна величина інтервалу варіювання визначається умовою адекватного опису об'єкта в області варіювання. Якщо при більших інтервалах варіювання математична модель неадекватна, то виникають систематичні помилки в визначенні коефіцієнтів, для зменшення яких необхідно звузити область варіювання. Однак зі зменшенням інтервалу варіювання з'являється цілий ряд нових труднощів: зростає відношення завади до корисного сигналу, що приводить до необхідності збільшувати кількість паралельних дослідів для виділення корисного сигналу на фоні "шуму", тобто зменшують ся абсолютні значення оцінок коефіцієнтів b_i, величини яких безпосередньо залежать від Δx_i (для рівняння з нормованими факторами z_i), та оцінки коефіцієнтів можуть стати статистично незначущими.

Для вибору інтервалу варіювання проводять попередні досліди. Інтервал варіювання можна вибрати рівним 0,05–0,3 від допустимого діапазону варіювання факторів, тобто область варіювання складає приблизно 10—60% від усього діапазону. Початкову точку варіювання (базову точку) вибирають якомога ближче до центра області факторного простору, в якій визначається математичний опис об'єкта (або області обмежень).

2.3 Дробовий факторний експеримент

В багатьох практичних задачах ідентифікації вплив взаємодій (добуток факторів) другого та вищих порядків відсутній або дуже малий. Крім того на перших етапах досліду часто необхідно отримати в першому наближенні лише лінійну апроксимацію рівняння зв'язку, що вивчається, при мінімальній кількості дослідів. Тому неефективно використовувати ПФЕ для оцінювання коефіцієнтів лише при лінійних членах та деяких парних добутках через реалізацію великої кількості варіантів варіювання 2^n, особливо при великій кількості факторів n. При лінійному рості кількості незалежних факторів кількість варіантів варіювання для ПФЕ зростає за степеневим законом, в результаті чого на перевірку гіпотези про адекватність залишається надто багато ступенів свободи.

Дробовим факторним експериментом (ДФЕ) називається експеримент, який реалізує частину (дробову репліку) повного факторного експерименту. ДФЕ дозволяє отримати, наприклад, лінійне наближення шуканої функціональної залежності $M\{y\} = \varphi(\bar{x})$ в деякому невеликому оточенні точки базового режиму при мінімумі дослідів.

Планування експерименту. Для розв'язання трифакторної (n=3) задачі регресії в лінійному наближенні можна задовольнитись чотирма варіантами варіювання, якщо в плануванні ПФЕ типу 2^2 додаток z_1z_2 прирівняти до третього незалежного фактору z_3. Таке планування, надане матрицею (табл. 2.2), дозволяє знайти вільний член b_0 та три оцінки коефіцієнтів регресії при лінійних членах b_1, b_2, b_3 (з чотирьох дослідів не можна отримати більше чотирьох оцінок коефіцієнтів регресії).

Використання ДФЕ завжди пов'язане зі змішуванням, тобто зі спільним оцінюванням декількох теоретичних коефіцієнтів математичної моделі. В даному випадку, якщо коефіцієнти регресії β_{il} при парних додатках відмінні від нуля, кожен із знайдених коефіцієнтів b_i служить оцінкою двох теоретичних коефіцієнтів регресії

$$b_0 \to \beta_0 + \beta_{123};\ \ b_1 \to \beta_1 + \beta_{23};\ \ b_2 \to \beta_2 + \beta_{13};\ \ b_3 \to \beta_3 + \beta_{12}.$$

Дійсно, вказані теоретичні коефіцієнти в такому плануванні не можуть бути оцінені окремо, оскільки стовпці МП для лінійних членів та парних додатків збігаються. Розглянутий план ДФЕ представляє собою половину плану ПФЕ типу 2^3 або планування типу $N=2^{3-1}$ (див. табл. 2.2).

Таблиця 2.2 – Матриця планування ДФЕ 2^{3-1}

g	z_0	z_1	z_2	z_3	z_1z_2	z_1z_3	z_2z_3	$z_1z_2z_3$
1	+1	–1	–1	+1	+1	–1	–1	+1
2	+1	+1	–1	–1	–1	–1	+1	+1
3	+1	–1	+1	–1	–1	+1	–1	+1
4	+1	+1	+1	+1	+1	+1	+1	+1

Для правильного планування ДФЕ необхідно використовувати всі отримані раніше відомості теоретичного та інтуїтивного характеру про об'єкт та виділити ті фактори та добутки факторів, вплив яких на відгук є значним. При цьому змішування потрібно проводити так, щоб лінійні коефіцієнти $\beta_0, \beta_1, \ldots, \beta_n$ були змішані з коефіцієнтами при взаємодії найвищого ступеня (оскільки звичайно вони в моделі відсутні), чи при тих взаємодіях, при яких апріорі відомо, що вони не впливають на відгук. Отже, недопустимо вільне розбиття плану ПФЕ типу 2^3 на дві частини для виділення напіврепліки типу 2^{3-1}.

При великій кількості n факторів для отримання лінійного наближення можна побудувати дробові репліки високого ступеня дробовості. Так, при $n=7$ можна скласти дробову репліку на основі ПФЕ типу 2^3, прирівнявши чотири із семи факторів до взаємодії трьох інших факторів: парних та потрійних. Будемо позначати тип дробової репліки записом 2^{n-p}, якщо p факторів прирівняні до добутків інших $n-p$ факторів.

План ДФЕ можна побудувати, прирівнюючи фактори різним взаємодіям (парним, потрійним тощо), зрозуміло що, при цьому змінюється система спільних оцінок теоретичних коефіцієнтів. Для отримання системи спільних оцінок і аналізу роздільної здатності дробових реплік зручно користуватись поняттями генерувального і визначального співвідношень. Генерувальне співвідношення служить для побудви дробової репліки. Так, в розглянутому плануванні ми задавали напіврепліку плану ПФЕ типу 2^3 за допомогою генерувального співвідношення $z_3=z_1z_2$.

Визначальним співвідношенням називається співвідношення, яке задає елементи першого стовпця матриці планування для фіктивної змінної (всі вони завжди дорівнюють 1). Вираз визначального співвідношення у випадку, який розглядається, отримується множенням лівої і правої частин приведеного генерувального співвідношення на z_3, тобто $1= z_1z_2z_3$, оскільки завжди $z_i^2 = 1$.

Знання визначального співвідношення дозволяє знайти всю систему спільних оцінок без вивчення матриці планування ДФЕ. Співвідношення, що задають ці оцінки, можна знайти послідовним множенням незалежних факторів на визначальне співвідношення:

$$z_0 = z_1z_2z_3;\quad z_1 = z_2z_3;\quad z_2 = z_1z_3;\quad z_3 = z_1z_2;$$

звідси легко знаходяться змішувані теоретичні коефіцієнти регресії та їх оцінки:

$$b_0 \to \beta_0 + \beta_{123};\quad b_1 \to \beta_1 + \beta_{23};\quad b_2 \to \beta_2 + \beta_{13};\quad b_3 \to \beta_3 + \beta_{12}.$$

Якщо апріорі можна прийняти, що коефіцієнти при всіх парних та потрійній взаємодіях рівні нулю, то реалізація цієї напіврепліки дозволить отримати роздільні оцінки для всіх чотирьох лінійних коефіцієнтів регресії. Роздільна здатність напівреплік визначається їх генерувальними співвідношеннями. Роздільна здатність тим вища, чим вищий порядок взаємодії, з коефіцієнтами яких змішані лінійні коефіцієнти. Вона збільшується для головних напівреплік з збільшенням кількості незалежних факторів.

Для чвертьрепліки в п'ятифакторному плануванні типу 2^{5-2} повинні бути задані два генерувальних співвідношення, наприклад $z_4 = z_1z_2z_3$; $z_5 = z_1z_2$, причому вважаємо $\beta_{123}=0$, тобто x_1, x_2, x_3 всі разом не взаємодіють, і $\beta_{12}=0$, тобто x_1 та x_2 також не взаємодіють. Визначальні співвідношення для цієї репліки, згідно з приведеними вище правилами, мають вигляд $1=z_1z_2z_3z_4$; $1=z_1z_2z_5$. Якщо для дробової репліки мають місце два чи більше визначальних співвідношення, то їх перемножують між собою, використовуючи всі можливі комбінації. У випадку, що розглядаємо, існує одна нова комбінація $1=z_3z_4z_5$.

Узагальнене визначальне співвідношення, що будується на основі всіх отриманих визначальних співвідношень, повністю характеризує роздільну здатність реплік високого ступеня дробовості. Так , в даному випадку

$$1 = z_1z_2z_3z_4 = z_1z_2z_5 = z_3z_4z_5.$$

Спільні оцінки тут визначаються допоміжними співвідношеннями:

$$z_0=z_1z_2z_3z_4=z_1z_2z_5=z_3z_4z_5,\qquad z_4=z_1z_2z_3=z_1z_2z_4z_5=z_3z_5,$$
$$z_1=z_2z_3z_4=z_2z_5=z_1z_3z_4z_5,\qquad z_5=z_1z_2z_3z_4z_5=z_1z_2=z_3z_4,$$
$$z_2=z_1z_3z_4=z_1z_5=z_2z_3z_4z_5,\qquad z_1z_3=z_2z_4=z_2z_3z_5=z_1z_4z_5,$$
$$z_3=z_1z_2z_4=z_1z_2z_3z_5=z_4z_5,\qquad z_1z_4=z_2z_3=z_2z_4z_5=z_1z_3z_5.$$

Ці допоміжні співвідношення дозволяють встановити, які стовпці МП будуть лінійно залежними і, отже, спільна оцінка таких теоретичних коефіцієнтів є той чи інший вибірковий коефіцієнт регресії:

$$b_0 \to \beta_0+\beta_{1234}+\beta_{125}+\beta_{345}, \qquad b_4 \to \beta_4+\beta_{123}+\beta_{1245}+\beta_{35},$$
$$b_1 \to \beta_1+\beta_{234}+\beta_{25}+\beta_{1345}, \qquad b_5 \to \beta_5+\beta_{12345}+\beta_{12}+\beta_{34},$$
$$b_2 \to \beta_2+\beta_{134}+\beta_{15}+\beta_{2345}, \qquad b_{13} \to \beta_{13}+\beta_{24}+\beta_{235}+\beta_{145},$$
$$b_3 \to \beta_3+\beta_{124}+\beta_{1235}+\beta_{45}, \qquad b_{14} \to \beta_{14}+\beta_{23}+\beta_{245}+\beta_{135}.$$

Роздільна здатність цієї чвертьрепліки невисока, оскільки всі теоретичні лінійні коефіцієнти регресії змішані з коефіцієнтами при парних взаємодіях. Слід мати на увазі, що план ДФЕ завжди можна доповнити до плану ПФЕ дробовими репліками, яких не вистачає. В даному прикладі для решти трьох чвертьреплік генерувальні співвідношення запишуться у вигляді:

$$\begin{cases} z_4 = z_1 z_2 z_3; \\ z_5 = -z_1 z_2; \end{cases} \quad \begin{cases} z_4 = -z_1 z_2 z_3; \\ z_5 = z_1 z_2; \end{cases} \quad \begin{cases} z_4 = -z_1 z_2 z_3; \\ z_5 = -z_1 z_2; \end{cases}$$

а узагальнювальні визначальні співвідношення – у вигляді:

$$1=z_1z_2z_3z_4=-z_1z_2z_5=-z_3z_4z_5,$$
$$1=-z_1z_2z_3z_4=z_1z_2z_5=-z_3z_4z_5,$$
$$1=-z_1z_2z_3z_4=-z_1z_2z_5=z_3z_4z_5.$$

Здійснення цих додаткових чвертьреплік означає реалізацію ПФЕ в цілому і, отже, окреме оцінювання всіх теоретичних коефіцієнтів регресії (якщо апріорі відомо, що $\beta_{ii}=0$, $\beta_{iii}=0$).

Проведення експерименту на об'єкті дослідження. Реалізація плану ДФЕ нічим не відрізняється від реалізації плану ПФЕ.

Перевірка відтворюваності експерименту. Перевірку однорідності оцінок дисперсії відгуку в різних точках факторного простору проводять в повній відповідності з методикою, викладеною для ПФЕ, різниця полягає лише в кількості точок плану.

Отримання математичної моделі об'єкта. Процедура визначення оцінок коефіцієнтів регресії та перевірки їх значущості повністю збігається з процедурою, що використовується при дослідженні об'єкта методом ПФЕ.

Перевірка адекватності математичного опису. Адекватність математичного опису функції відклику перевіряють тими ж методами, що і для ПФЕ.

ЗАВДАННЯ ТА ПОРЯДОК ВИКОНАННЯ РОБОТИ

1. Побудувати матриці планування ПФЕ для двох, трьох і ДФЕ для чотирьох незалежних факторів, які варіюють на двох рівнях.
2. Задати модель об'єкта №3 з двома факторами x_1, x_2.

3. Провести експеримент у відповідності зі створеним планом ПФЕ типу 2^2 з центром в точці $\bar{x}_0$ з координатами $x_{10} = 5{,}0$; $x_{20} = 6{,}0$ та інтервалами зміни $\Delta x_1 = \Delta x_2 = 0{,}2$ при середній квадратичній помилці відтворюваності експерименту $\sigma_y = 5$. Провести три серії паралельних дослідів ($m = 3$).
4. Задати модель об'єкта №4 з трьома факторами x_1, x_2, x_3.
5. Провести експеримент у відповідності зі складеним планом ПФЕ типу 2^3 з центром в точці $\bar{x}_0$ з координатами $x_{10} = 5{,}0$; $x_{20} = 6{,}0$; $x_{30} = 5{,}0$ та інтервалами варіювання $\Delta x_1 = \Delta x_2 = \Delta x_3 = 0{,}2$ при середній квадратичній помилці відтворюваності експерименту $\sigma_y = 5$. Провести три серії паралельних дослідів ($m = 3$).
6. Обробити результати спостережень експериментів п. 3 і 5:

а) провести перевірку відтворюваності експерименту;

б) розрахувати оцінку коефіцієнтів регресії, визначити дисперсії оцінок коефіцієнтів, перевірити їх значимість;

в) перевірити адекватність отриманого математичного опису.

7. Знайти визначальні співвідношення напівреплік планування типу 2^{4-1} з генерувальним співвідношеннями:

1) $z_4 = z_1 z_2 z_3$; 2) $z_4 = -z_1 z_2 z_3$; 3) $z_4 = z_1 z_2$; 4) $z_4 = -z_1 z_2$;

5) $z_4 = z_1 z_3$; 6) $z_4 = -z_1 z_3$; 7) $z_4 = z_2 z_3$; 8) $z_4 = -z_2 z_3$.

8. Вибрати з цих напівреплік одну для реалізації ДФЕ, якщо апріорі відомо, що на відгук можуть впливати лише три парних взаємодії $x_1 x_2$, $x_2 x_3$, $x_2 x_4$ і лінійні члени. Вибір виконувати з умови одержання незмішаних оцінок лінійних коефіцієнтів і коефіцієнтів трьох означених парних взаємодій. Побудувати МП вибраної напівреплiки.
9. Задати модель об'єкта №5 із чотирма факторами x_1, x_2, x_3, x_4.
10. Провести експеримент у відповідності з планом вибраної напівреплiки типу 2^{4-1} з центром в точці $\bar{x}_0$ з координатами $x_{10} = 5{,}0$; $x_{20} = 6{,}0$; $x_{30} = 5{,}0$; $x_{40} = 4{,}0$ та інтервалами варіювання $\Delta x_1 = \Delta x_2 = \Delta x_3 = 0{,}2$; $\Delta x_4 = 0{,}3$ при середній квадратичній помилці відтворюваності експерименту $\sigma_y = 5$. Провести три серії паралельних дослідів ($m = 3$).
11. Обробити результати спостережень експерименту (п. 9) у відповідності з програмою п. 6.
12. Дати графічну інтерпретацію рішення двофакторної задачі регресії, побудувавши на площині двох факторів x_1, x_2 лінії постійних рівнів відгуку ($\hat{y} = const$) за допомогою отриманого

рівняння регресії, і нанести в точках варіантів варіювання середні дослідницькі значення відгуку $\bar{y}_g$.

ФОРМУЛИ ДЛЯ РОЗРАХУНКУ

1. $\bar{y}_g = \dfrac{1}{m}\sum\limits_{k=1}^{m} y_{gk}$.

2. 2. $s_g^2 = \dfrac{1}{m-1}\sum\limits_{k=1}^{m}\left(y_{gk} - \bar{y}_g\right)^2$ з числом ступенів свободи $\nu_{1ВІДТ} = m - 1$.

3. $G = \dfrac{\max\limits_g\left\{s_g^2\right\}}{\sum\limits_{g=1}^{N} s_g^2} < G_p\left(\nu_{1ВІДТ}; \nu_{2ВІДТ}\right)$ з числом ступенів свободи $\nu_{1ВІДТ} = m - 1$ і $\nu_{2ВІДТ} = N$.

4. $b_0 = \dfrac{1}{N}\sum\limits_{g=1}^{N} z_{0g}\bar{y}_g$. 5. $b_i = \dfrac{1}{N}\sum\limits_{g=1}^{N} z_{ig}\bar{y}_g$ $(i = 1, 2, \dots, n)$.

6. $b_{il} = \dfrac{1}{N}\sum\limits_{g=1}^{N} z_{ig}z_{lg}\bar{y}_g$ $(i; l = 1, 2, \dots, n; i \neq l)$.

7. $s_{ВІДТ}^2\{y\} = \dfrac{1}{N}\sum\limits_{g=1}^{N} s_g^2$ з числом ступенів свободи $\nu_{ЗН} = N(m-1)$.

8. $s^2\{b\} = \dfrac{1}{Nm} s_{ВІДТ}^2\{y\}$ з числом ступенів свободи $\nu_{ЗН} = N(m-1)$.

9. $t = |b|/s\{b\} > t_{кр}(\nu_{ЗН})$ з числом ступенів свободи $\nu_{ЗН} = N(m-1)$.

10. $s_{АД}^2 = \dfrac{m}{N-d}\sum\limits_{g=1}^{N}\left(\bar{y}_g - y_g\right)^2$ з числом ступенів свободи $\nu_{1АД} = N - d$.

11. $F = s_A^2 / s_{ВІДТ}^2\{y\} < F_p(\nu_{1A}; \nu_{2A})$ з числом ступенів свободи $\nu_{1АД} = N - d$ і $\nu_{2АД} = \nu_{ЗН} = N(m-1)$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що називається повним і дробовим факторним експериментом?
2. Як вибираються фактори планування, їх основні (базові) рівні й інтервали варіювання?
3. Зазначити порядок проведення експерименту методом ПФЕ?
4. Як складається матриця планування ПФЕ?
5. Як перевірити відтворюваність варіантів варіювання ПФЕ?
6. При яких умовах не дотримується вимога відтворюваності експерименту і що варто зробити в цьому випадку?
7. Як перевірити значимість оцінок коефіцієнтів регресії?
8. При яких умовах оцінки коефіцієнтів регресії є незначимими і як ці умови усунути?
9. Як перевірити адекватність математичної моделі?
10. При яких умовах не дотримується вимога адекватності математичної моделі і що варто зробити в цьому випадку?
11. Що таке генерувальне співвідношення і як воно вибирається?
12. Що таке визначальне співвідношення і як з його допомогою складається система спільних оцінок?
13. Від чого залежить роздільна здатність дробових реплік?

Тема 3 МЕТОД ВИПАДКОВОГО БАЛАНСУ

Мета – *ознайомлення з основними ідеями і процедурою застосування методу випадкового балансу, призначеного для виділення найбільш істотних вхідних змінних серед великого числа лінійних факторів і їхніх парних взаємодій у багатофакторному об'єкті.*

КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

3.1 Основні ідеї та передумови

При оптимізації багатофакторного об'єкта основним етапом є отримання математичної моделі, що адекватно описує статичний об'єкт у діапазоні зміни його вхідних змінних (факторів). При цьому природно прагнути до того, щоб математичний опис був якомога більш простим при максимумі подібності, особливо при розробці способів та систем оптимального управління, коли важливо досягнути або підтримувати глобальний, а не локальний або частковий екстремум. Проте вирішення цієї задачі в реальних умовах зазвичай пов'язане з серйозними утрудненнями, які викликані дуже великою кількістю змінних, які в тій чи іншій мірі впливають на об'єкт. Методика регресивного аналізу основана на припущенні, що враховані всі або, у крайньому разі, всі істотні фактори, інакше отримана математична модель стане неадекватною. Залучення всієї множини змінних до складання математичного опису може вимагати непомірного об'єму експериментальної та обчислювальної роботи, що найчастіше не може бути виконана через технологічні, економічні та інші обмеження. Виникає необхідність попереднього відсіву неістотних змінних та виділення тих вхідних впливів, які найбільш відчутно впливають на цільову функцію.

Якщо число всіх можливих факторів, що впливають на об'єкт, не перевищує шести–семи, то для попереднього вивчення об'єкта можна використати методи дробового або повного факторного експерименту. Проте при великому числі розглянутих факторів методи ПФЕ і навіть ДФЕ, призначені для ретельного вивчення поверхні відгуку, виявляються занадто громіздкими та трудомісткими для постановки дослідів відсіювання. Якщо накласти деякі обмеження при самій постановці задачі, то можна знизити кількість таких дослідів. Наприклад, коли діють тільки лінійні змінні, використовують насичені плани ДФЕ, при яких всі $N-1$ ступенів свободи (N – кількість рядків МП, що відповідає кількості точок факторного простору, в яких ставляться досліди) використовуються для оцінки коефіцієнтів при лінійних членах рівняння регресії. Метод випадкового балансу (МВБ) використовує перенасичені плани, при яких виконується умова

$$u+1>N, \qquad (3.1)$$

де u – загальна кількість досліджуваних змінних, в тому числі l лінійних факторів і $u-l$ парних взаємодій, а l представляє постійний член.

Важливою теоретичною передумовою МВБ є апріорне знання того, що з усієї сукупності розглянутих змінних тільки невелика їх кількість (наприклад, 10–15%) є дійсно суттєвими, інші ж можуть бути віднесені до «шумового поля». Під «шумовим полем» розуміють випадкові перешкоди ε, про які нічого невідомо, та малозначимі і незначимі змінні (лінійні та парні взаємодії), які недоцільно контролювати. Постулюється, що для успішного застосування МВБ розглянуті змінні розташовані в порядку спадання внесків в загальну дисперсію функції відгуку і повинні утворити криву, що швидко загасає (рис. 7). Внески істотних факторів повинні набагато (мінімум на порядок) перевищувати похибку вимірювання, що визначається шумовим полем. Крім того, передбачається: що об'єкт управляється; що між окремими складовими дисперсії функції відгуку і вхідними змінними може бути встановлена відповідність; що досліди відтворені, а окремі виміри випадкові і незалежні один від одного (виконуються теоретичні передумови регресійного аналізу).

Основна ідея методу полягає в тому, що замість дробових реплік, що є систематичними ортогональними вибірками з ПФЕ, беруться випадкові вибірки. Тоді вектори–стовпці матриці планування можна вважати не корельованими або слабко корельованими один з одним. Спільні оцінки стають змішаними випадковим чином. З'являється можливість виділити і незалежно оцінити всі змінні, що домінують.

Формально нерівність (3.1) означає, що число ступенів свободи $\nu<0$. При цій умові, зрозуміло, неможливо одержати незалежні кількісні оцінки коефіцієнтів рівняння регресії для всіх розглянутих змінних. Але цього і не потрібно при проведенні дослідів відсіювання, завданням яких є виділення істотних змінних і віднесення неістотних або мало істотних змінних до "шумового поля". Тоді виділені істотні змінні (лінійні члени і парні взаємодії) можуть бути незалежно оцінені і кількісно, якщо їхнє число h підпорядковується умові

$$h+1<N, \qquad (3.2)$$

причому це оцінювання виконується, очевидно, із тим більшою похибкою, чим більшою є частка дисперсії функції відгуку, що викликається "шумовим полем". При використанні МВБ математичну модель складного об'єкта

$$\hat{y} = \hat{a}_0 + \sum_{i=1}^{h} \hat{a}_i z_i + \sum_{r=h+1}^{n} \hat{a}_r z_r + \varepsilon \qquad (3.3)$$

розщеплюють на частини, де h – число значимих змінних; $u–h$ – число незначимих змінних; ε – випадкова складова похибки. Під z_i, z_r розуміються нормовані лінійні фактори і парні взаємодії.

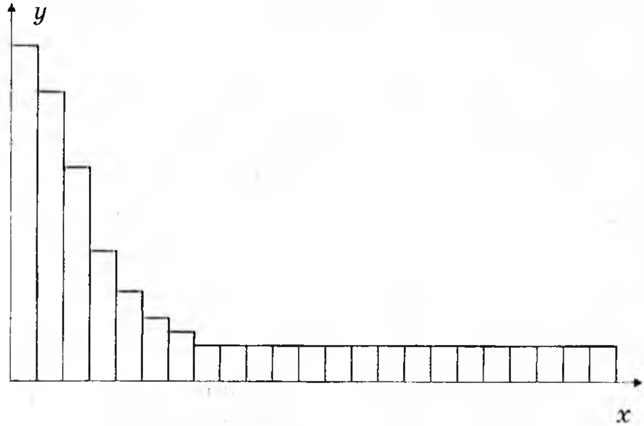


Рисунок 7 – Залежність змінних від їх внесків: x – змінні, y – внески змінних

Із сказаного випливає, що МВБ має меншу чутливість, ніж ПФЕ і ДФЕ. Під чутливістю методу звичайно розуміється спроможність виділяти коефіцієнти a_i рівняння регресії, що значно відрізняється від нуля (тобто спроможність відкидати нуль–гіпотезу H_0: a_i=0. Зате він має більшу роздільну здатність у сприятливих умовах, при однаковому числі дослідів МВБ дозволяє незалежно виділити істотні змінні серед набагато більшого числа змінних, що переглядались, ніж при ДФЕ і тим більше ПФЕ. Метод варто застосовувати при вивченні більше восьми–десяти факторів, якщо експерименти недорогі і якщо явно відомо, що лише деякі змінні є істотними. Якщо ж майже всі або хоча б більшість змінних, що розглядаються, визначають процес і їх небагато, то краще застосовувати ДФЕ з підходящими визначальними контрастами.

3.2 Побудова матриці планування

Побудову матриці планування для проведення дослідів відсіювання виконують на основі передумови, що досліджувані фактори повинні бути змішані випадковим чином. Всі лінійні фактори z_i ($i = 1, 2, ..., n$) розбивають на групи, при цьому прагнуть фактори, що явно взаємодіють, включити в одну групу. Якщо ж немає апріорних відомостей про фізику процесу, то розбиття факторів по групах можна робити формально, з використанням таблиці випадкових чисел. У цьому випадку попередньо складають пронумерований список факторів, а потім за допомогою таблиці випадкових чисел кожному чиннику привласнюють свій випадковий порядковий номер, після чого, наприклад, фактори під випадковими номерами з 1 по 4 об'єднують у першу групу, фактори під номерами з 5 по 8 – у другу і т.д. Потім для кожної групи складають МП на основі ПФЕ і ДФЕ. Всі групові МП повинні мати однакову кількість рядків, щоб їх можна було сортувати. Число N рядків кожної групової матриці повинно задовольняти умову (3.1) і рівність

$$N=2^p. \qquad (3.4)$$

Причому p вибирається звичайно найменшим, при якому виконується нерівність $N>n$, де n – число лінійних факторів. План експерименту відсіювання створюють шляхом стикування групових МП випадковим змішуванням їхніх рядків.

Пояснимо це на прикладі: Нехай потрібно дослідити шість факторів z_1, z_2, z_3, z_4, z_5, z_6 і 15 їхніх парних взаємодій і, за допомогою МВБ, виділити самі істотні. Розіб'ємо 6 факторів на дві групи: 1 – z_1, z_2, z_3; 2 – z_4, z_5, z_6. Тоді групі 1 відповідає матриця ПФЕ типу 2^3 (табл. 3.1), а групі 2– аналогічна (табл. 3.2).

Загальна МП для МВБ складається у вигляді табл.3.3 шляхом порядкового сортування табл. 3.1 і 3.2 після рандомізації їхніх рядків за допомогою таблиці випадкових чисел (див. додаток А). Початок відліку в таблиці випадкових чисел повинен вибиратися довільно. Випадкові числа вибираються “без повторення”, тобто одне число береться тільки один раз. У цьому випадку отримана МП є нормальною (симетричною), тобто кожний із факторів планується однакове число разів як на рівні −1, так і на рівні +1. Це підвищує точність розрахунків. Загальна МП у даному прикладі може мати вигляд табл. 3.3.

Порядок номерів g_1 і g_2 обраний із таблиці випадкових чисел, причому спочатку обрана серія g_1, а початок вибору серії g_2 повинен був служити кінцем вибору серії g_1. При упорядкуванні загальної МП для МВБ необхідно, щоб максимальне число її векторів–стовпців було ортогональним, причому в ортогональну частину прагнуть включити найбільш істотні фактори, якщо їх можна оцінити апріорно. Якщо вдасться вгадати всі істотні фактори і включити їх в одну групу, то це означає, що для них одночасно виявиться здійсненним

ПФЕ. Результати МВБ аналізуються за допомогою діаграм розсіювання або за допомогою вибіркових ортогональних МП.

Таблиця 3.1

Номер рядка (g_1)	z_1	z_2	z_3
1	−1	−1	−1
2	+1	−1	−1
3	−1	+1	−1
4	+1	+1	−1
5	−1	−1	+1
6	+1	−1	+1
7	−1	+1	+1
8	+1	+1	+1

Таблиця 3.2

Номер рядка (g_2)	z_4	z_5	z_6
1	−1	−1	−1
2	+1	−1	−1
3	−1	+1	−1
4	+1	+1	−1
5	−1	−1	+1
6	+1	−1	+1
7	−1	+1	+1
8	+1	+1	+1

Таблиця 3.3

Номер рядка загальної МП (g)	Номер рядка табл. 3.1 (g_1)	z_1	z_2	z_3	Номер рядка табл. 3.2 (g_2)	z_4	z_5	z_6
1	3	−1	+1	−1	4	+1	+1	−1
2	4	+1	+1	−1	7	−1	+1	+1
3	7	−1	+1	+1	8	+1	+1	+1
4	5	−1	−1	+1	5	−1	−1	+1
5	2	+1	−1	−1	3	−1	+1	−1
6	8	+1	+1	+1	6	+1	−1	+1
7	6	+1	−1	+1	2	+1	−1	−1
8	1	−1	−1	−1	1	−1	−1	−1

3.3 Побудова діаграм розсіювання

Побудову діаграми розсіювання пояснимо на прикладі. Нехай досліджується вплив шести факторів за планом, наведеним в табл. 3.3, а результати експерименту дані в табл. 3.4.

Таблиця 3.4

g	k	z_1	z_2	z_3	z_4	z_5	z_6	y_g	y^{I}_g
1	5	−1	+1	−1	+1	+1	−1	27	39,5
2	4	+1	+1	−1	−1	+1	+1	49	49
3	8	−1	+1	+1	+1	+1	+1	31	43,5
4	6	−1	−1	+1	−1	−1	+1	39	39
5	2	+1	−1	−1	−1	+1	−1	64	64
6	7	+1	+1	+1	+1	−1	+1	40	52,5
7	3	+1	−1	+1	+1	−1	−1	42	54,5
8	1	−1	−1	−1	−1	−1	−1	47	47

У графі k наведено випадковий порядок проведення дослідів по рядках МП (взагалі, в даному випадку досліди можна проводити і по порядку номерів g, оскільки загальна матриця, дякуючи випадковому підбору рядків групових матриць, вже рандомізована). Для зменшення впливу перешкод і одержання оцінки дисперсії відтворюванності можуть виконуватися паралельні досліди. Для кожного фактора проводиться своя ордината (рис. 8). зліва від неї відзначаються точками ті значення відгуку y_k, що відповідають положенню даного фактора на нижньому рівні варіювання $z_{ig}= -1$, а справа – для $z_{ig}= +1$(i – номер фактора; g – номер рядків МП). Потім знаходяться частинні медіани (в якості центру розподілення в МСБ використовують медіану, оскільки при асиметричних розподіленнях вона більш ефективна, ніж середнє арифметичне значення) – окремо для випадкового розсіювання точок зліва від ординати й окремо для точок справа. Правильність побудови діаграм розсіювання необхідно перевіряти підрахунком кількості N точок, кількості точок на кожному з рівнів, а також перевіркою, чи присутнє кожне y_g (g = 1, 2, ..., N) в діаграмах кожного z_i, тобто перевіркою по горизонталі. Нагадаємо, що вибірковою оцінкою медіани $y=Me\{y\}$ для дискретних даних у математичній статистиці називається таке значення випадкової величини y, по обох сторонах якого лежать рівні кількості випадкової величини y, незалежно від їхніх конкретних значень. Якщо кількість точок парна ($2N$), то медіана лежить посередині між N-ю і (N+1)-ю точками, якщо ж – непарна ($2N$+1), то медіаною є (N+1)-а точка. Різниця між медіаною справа $Me\{y\}\big|_{z_i=+1}$ і медіаною зліва $Me\{y\}\big|_{z_i=-1}$ (але не навпаки) називається внеском фактора z_i у цільову функцію y і позначається B_{z_i} (для взаємодії $B_{z_i}B_{z_r}$)

$$B_{z_i} = Me\{y\}\big|_{z_i=+1} - Me\{y\}\big|_{z_i=-1}\,. \qquad (3.5)$$

З рис. 8, що побудований за даними табл. 3.4, знаходимо:

B_{z_1} = 45,5 – 35,0 = 10,5; B_{z_2} = 35,5 – 44,5 = – 9,0;

B_{z_3} = 39,5 – 48,0 = – 8,5; B_{z_4} = 35,5 – 48,0 = – 12,5;

B_{z_5} = 40,0 – 41,0 = – 1,0; B_{z_6} = 39,5 – 44,5 = – 5,0.

Візуальне і чисельне порівняння внесків дає можливість виявити, що найбільш істотні змінні мають найбільші внески. Але в ряді випадків абсолютний розмір внесків не є достатнім критерієм найбільшої істотності змінних. Тому використовують також додаткові критерії: корельованність векторів–стовпців, наявність, число і характер точок, що виділяються. Пояснимо поняття “точки, що виділяються” на рис. 8 для фактора z_2 на рівні $z_2= -1$ є одна точка, розташована вище, ніж найвища точка на рівні $z_2= +1$, а на рівні $z_2= +1$ є дві точки, розташовані нижче, ніж найнижча точка на рівні $z_2= -1$. Таким чином, z_2 має три точки, що виділяються, z_1 і z_4 – по п'ять точок, а z_3, z_5 і z_6 – жодної, так що, незважаючи на приблизно однакові внески z_2 і z_3, чинник z_3 визнається несуттєвим у порівнянні з z_2 і тим більше з z_4 і z_1. Фактори z_3, z_5 і z_6 не мають точок, що виділяються тому, що в цих трьох факторах розмах точок на одному з рівнів перевищує розмах точок на іншому рівні, причому так, що на рівні, де розмах менше, немає точок розташованих ні вище, ні нижче. Необхідно враховувати і характер розташування точок, що виділяються: чи розташовуються вони відповідно до знака внеску або протилежно до нього.

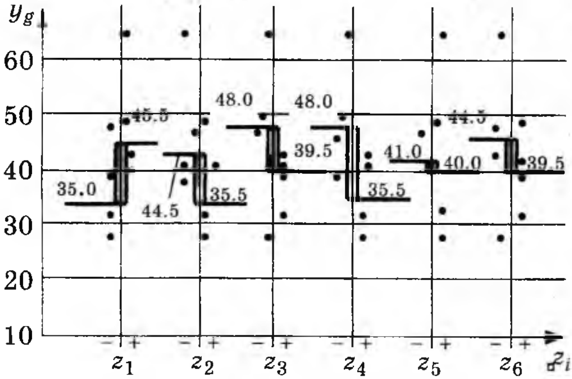


Рисунок 8 – Діаграма розсіювання

3.4 Послідовне виділення істотних змінних

Виділення найбільш істотних змінних і їхнє ранжування можна зробити двома способами: за допомогою внесків і ортогональних вибіркових МП. Щоб виключити вплив найбільш істотних факторів за допомогою внесків, їх треба ніби “стабілізувати” на одному з двох рівнів варіювання. У розглянутому прикладі (рис. 8) варто спочатку виділити z_4 ($B_{z_4} = -12.5$ і п'ять точок, що виділяються,). Стабілізація z_4 здійснюється вирахуванням внеску B_{z_4} зі своїм знаком із величини y_g у тих рядках табл. 3.4, де z_{4g}=+1 (при стабілізації z_4 на рівні $z_4 = -1$), тобто за допомогою формули

$$y_g^I = y_g - B_{zi}, \qquad (3.6)$$

де y_g – вихідне (дослідне) значення відгуку в g–у рядку вихідної МП (див. табл. 3.4); y_g^I – значення відгуку в g–у рядку після першого (індекс I) коригування; B_{z_i} – внесок найбільш істотного нормованого фактора. Значення y_g на іншому рівні варіювання (z_4=–1) залишається при цьому незмінним. Результати першого коригування представлені в останньому стовпці табл. 3.4. За новими даними y_g будують нову діаграму розсіювання і вже по ній визначають таку змінну, яка за рангом впливу має найбільший внесок, після чого описану процедуру повторюють. Очевидно, раніше виділені ("стабілізовані") найбільш істотні змінні на кожній наступній діаграмі розсіювання не підлягають вивченню: відповідні їм координати залишаються незаповненими. Якщо виявиться, що внески двох факторів приблизно рівні, то більш істотним визнають той, у діаграмі розсіювання якого значно більше точок, що виділяються, у верхній і нижній частинах діаграми, навіть якщо він має дещо менший за величиною внесок. Наприклад, із рис. 8 очевидно, що $B_{z_1} \approx B_{z_2}$, проте сумарна кількість точок, що виділяються, для z_1 дорівнює 2+3=5. Для z_2 є лише три точки, що виділяються. Тому z_1 приймається більш істотним і він виключається раніше. Процес виділення істотних змінних припиняють, коли виконана умова (3.2) або коли на черговій діаграмі розсіювання всі внески B_{z_i} стають приблизно однаково малими за абсолютною величиною і порівнянними за t-критерієм Стьюдента з подвоєною помилкою коефіцієнтів нормованого рівняння регресії. Методом випадкового балансу, як це випливає з співвідношень (3.2) і (3.3), можна одержати рівняння регресії

$$y = a_0 + a_1 \cdot z_1 + a_2 \cdot z_2 + \ldots + a_h \cdot z_h + \xi, \qquad (3.7)$$

де z_1, z_2, ..., z_h – виділені найбільш істотні нормовані лінійні фактори і їхні взаємодії, h – задовольняє умову (3.2); ξ – усе "шумове поле". Описаний порядок виділення і ранжування істотних змінних за допомогою діаграм розсіювання і внесків являє собою, очевидно, трудомістку процедуру, що потребує тим більшої уваги, чим більше загальне число і змінних, і взаємодій.

Цю трудомісткість можна знизити, якщо застосувати ортогональні вибіркові МП для 2–3 найбільш істотних факторів z_i, оцінених за першою діаграмою розсіювання, побудованою за вихідною МП. Побудову ортогональної вибіркової МП для двох факторів розглянемо на прикладі табл. 3.4 і рис. 8. Візуальне оцінювання і розрахунок внесків із залученням додаткових критеріїв показали, що найістотнішими факторами є z_4 і z_1, що мають і найбільше число точок, що виділяються (по п'ять). Побудуємо ортогональну МП для n=2 (табл. 3.5), у якій прийняті такі позначення: $g_{виб}$ – номер рядка вибіркової ортогональної МП типу $N=2^2$. Таким чином, у кожному рядку МП (табл. 3.5) виявилося по два паралельних досліди (y_{g1} і y_{g2}), порядкові середні значення відгуку дані в графі $y_{gвиб}$. Табл. 3.5 дозволяє розрахувати оцінки коефіцієнтів нормованого рівняння регресії за формулою (2.23), а також зробити їхнє статистичне оцінювання. У графі $s^2{}_{gвиб}$ приведені незміщені оцінки порядкових дисперсій, їхня однорідність перевіряється за допомогою G-критерію Кохрена, якщо в кожному рядку вибіркової МП є однакове число m паралельних дослідів. В останній графі табл. 3.5 підраховані сума порядкових дисперсій ($\Sigma s^2{}_{gвиб}$=154,5) та розрахункове значення критерію Кохрена (G=0,73). Критичне значення $G_{кр}$ знайдено з додатка Б для числа ступенів свободи чисельника ν_1=1 та знаменника ν_2=4 (порівнюються чотири дисперсії) при q=0.05. Оскільки $G<G_{кр}$, то можна зробити висновок про однорідність оцінок дисперсій.

Таблиця 3.5

$g_{виб}$	z_4	z_1	y_{g1}	y_{g2}	$y_{gвиб}$	$s^2{}_{gвиб}$	Перевірка відтворюваності
1	−1	−1	39	47	43	32	$\sum s^2{}_{gвиб}$=154,5
2	+1	−1	27	31	29	8	G=0,73
3	−1	+1	49	64	56.5	112.5	$G_{кр}$=0,906
4	+1	+1	40	42	41	2	

Якщо оцінки коефіцієнтів $a_{4виб}$ та $a_{1виб}$ стануть статично значимими, то приступають до виділення найбільш істотних факторів z_4 та z_1. Тому вихідну табл. 3.4 коректують з допомогою $a_{4виб}$ та $a_{1виб}$ за формулою, аналогічною формулі (3.6), стабілізуючи z_4 та z_1 на рівнях z_4= −1 та z_1= −1

$$y_g{}^{\mathrm{I}}=y_g - 2a_{4виб}-2a_{1виб}; \qquad (3.8)$$

з цією метою в тих g–рядках, де z_{4g}= +1 та z_{1g}= +1, з вихідних значень відгуку y_g віднімають подвоєні значення коефіцієнтів $a_{4виб}$ та $a_{1виб}$ із своїми знаками ($a_{4виб}$= − 7.49, $a_{1виб}$= + 6.4). Віднімати подвоєні коефіцієнти рівняння регресії слід тому, що a_i характеризує середню зміну відгуку y при переході нормованого фактора з базового рівня z_i= 0 на рівень z_i= +1, в той час як внесок B_{zi} характеризує усереднену зміну відгуку y при переході z_i з рівня $z_i= -1$ на рівень $z_i= +1$.

В результаті коректування отримаємо таблицю (табл. 3.6). За даними цієї таблиці можна побудувати нову діаграму розсіювання (рис. 9), з якої видно, що: для z_2 зменшилась абсолютна величина внеску B_{z_i}; для z_3 зменшилась величина внеску та не змінився характер діаграми (не з'явилось жодної точки, що виділяється); для z_5 характер діаграми не змінився, проте змінився знак вкладу, що буває у неістотних факторів; для z_6 абсолютна величина вкладу не змінилась, проте тільки у цього фактора з усіх, що залишились, змінився характер діаграми розсіювання (з'явились чотири точки, що виділяються). Розглядаючи відкоректовану діаграму розсіювання бачимо, що на другому етапі виділення найбільш істотних змінних потрібно виключити вплив z_2 та z_6, які мають найбільші внески та по три –чотири точки, що виділяються. Для такого виділення можна використовувати або внески B_{z_6} та B_{z_2}, або коефіцієнти $a_{6виб}$ та $a_{2виб}$, знайдені з ортогональної вибіркової МП, складеної для z_2 та z_6 аналогічно табл. 3.5. На чергових діаграмах розсіювання ординати для виділених найбільш істотних змінних залишають незаповненими. Відмітимо, що дисперсія точок в діаграмах розсіювання після виділення найбільш істотних факторів (рис. 9) відчутно зменшилась в порівнянні з початковою картиною (рис. 8). При великій кількості n основних факторів можна побудувати ортогональну вибіркову МП типу 2^{4-i} або 2^{3-i}, якщо вдається вибрати з вихідної МП достатню кількість відповідних рядків.

Таблиця 3.6

g	z_1	z_2	z_3	z_4	z_5	z_6	y_g	y^{I}_g
1	−1	+1	−1	−1	+1	−1	27	41,8
2	−1	+1	−1	−1	+1	+1	49	36,2
3	−1	+1	+1	−1	+1	+1	31	45,8
4	−1	−1	+1	−1	−1	+1	39	39
5	−1	−1	−1	−1	+1	−1	64	51,2
6	−1	+1	+1	−1	−1	+1	40	42
7	−1	−1	+1	−1	−1	−1	42	44
8	−1	−1	−1	−1	−1	−1	47	47

Розглянемо операції з парними взаємодіями. В табл. 3.4 та на рис. 8 ми мали діло тільки з лінійними факторами. Але табл. 3.4 можна доповнити вектор-стовпцями парних взаємодій, знаки яких отримуються простим перемноженням знаків факторів у відповідних рядках МП (табл. 3.4). Діаграму розсіювання (рис. 8) можна доповнити ординатами для парних взаємодій та аналізувати їх описаним способом нарівні з лінійними факторами. Але число взаємодій швидко зростає із збільшенням числа n факторів. Передбачається, що найбільш істотну взаємодію, наприклад $z_i z_r$, порівняно з іншими взаємодіями повинно мати і найбільше число точок, що виділяються, на обох рівнях. Але для того щоб добуток $z_{ig} z_{rg}$ був від'ємним, g-і точки відгуку y_g факторів z_i та z_r повинні бути на різних рівнях, а для того щоб він був додатним – на однакових рівнях. Очевидно, найбільше число точок, що виділяються (тобто взагалі і найбільший вклад $B_{z_i z_r}$), виявиться в тому випадку, якщо z_i та z_r в одній із частин діаграми розсіювання, наприклад у верхній, мають однаковий за рівнями розкид відповідних g-х точок, а в іншій частині (нижній) – протилежний, ніби дзеркально відображений розкид за рівнями g–х точок. Це видно з рис. 10, де показані тільки верхні та нижні частини діаграми розсіювання, побудовані за абстрактними даними, а середні частини зображені пунктирними лініями. Інколи навіть порівняно слабкі лінійні фактори можуть породити досить істотні парні взаємодії.

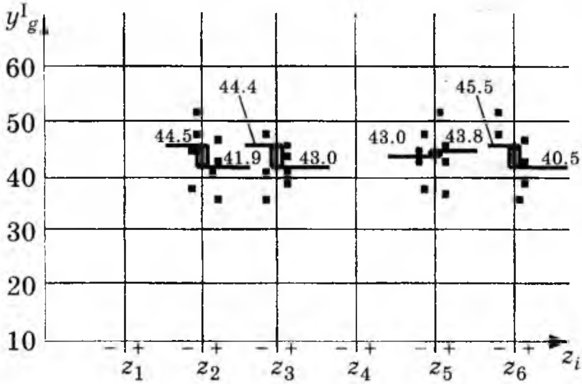


Рисунок 9 – Діаграма розсіювання

3.5 Обчислення оцінок коефіцієнтів та статистична оцінка результатів

Оцінки коефіцієнтів нормованого рівняння регресії можна отримати по вкладах:

$$a_i = 0.5\, B_{z_i}\ ;\ a_{ir} = 0.5\, B_{z_i z_r}\ . \qquad (3.9)$$

Розрахування вкладів для кожного менш істотного фактора виконують на основі діаграми розсіювання, побудованої за відкоректованими даними, після виділення більш істотних змінних. Критерій для припинення виділення найбільш істотних факторів при використанні вкладів будується на основі наближеної оцінки дисперсії коефіцієнтів a_i

$$B_{z_i} = B_{zкр} = 2t_{кр}s\{a_i\}, \tag{3.10}$$

$$s^2\{a_i\} = \frac{1}{Nm}s^2_{відт}\{y\}, \tag{3.11}$$

причому N – загальне число рядків вихідної МП для МВБ, m – число паралельних дослідів в ній; осереднена оцінка дисперсії

$$s^2_{відт}\{y\} = \frac{1}{N}\sum_{g=1}^{N}s^2_g\{y\}, \tag{3.12}$$

$s^2_g\{y\}$ – порядкова дисперсія в g-у рядку вихідної МП для МВБ; число ступенів свободи, за якими з додатка В вибирають $t_{кр} = t_{табл}\{\nu; q\}$, складає

$$\nu = N(m-1). \tag{3.13}$$

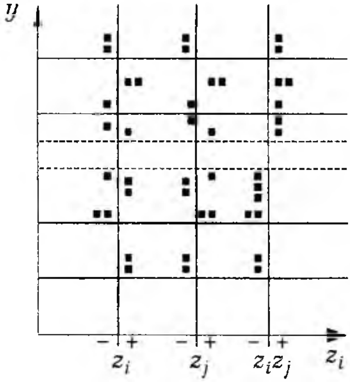


Рисунок 10 – Діаграма розсіювання

При використанні вкладів розрахунок та статистичне оцінювання B_{z_i} або a_i проводять послідовно, причому розрахунок та оцінювання істотності чергового фактора виконують за відкоректованими даними після виділення попереднього, більш істотного фактора.

При використанні вибіркових ортогональних матриць (ВОМ) обчислення оцінок $a_{iвиб}$ коефіцієнтів виконують за формулою, що враховує число $N_{виб}$ рядків ВОМ

$$a_{i_{виб}} = \frac{1}{N_{виб}}\sum_{g_{виб}=1}^{N_{виб}} z_{ig_{виб}}\bar{y}_{g_{виб}}. \tag{3.14}$$

Критичне значення вибіркової оцінки $a_{iвиб.кр}$ знаходять за допомогою формул:

$$a_{i_{виб.кр}} = t_{кр}s\{a_{i_{виб}}\}, \tag{3.15}$$

$$s^2\{a_{i_{виб}}\} = \frac{1}{N_{виб}m_{виб}m}s^2_{відт.виб}\{y\}, \tag{3.16}$$

$$s^2_{відт.виб}\{y\} = \frac{1}{N_{виб}} \sum_{g_{виб}=1}^{N_{виб}} s^2_{g_{виб}}\{y\}, \qquad (3.17)$$

$$s^2_{g_{виб}}\{y\} = \frac{1}{m_{виб}-1} \sum_{k=1}^{m_{виб}} (y_{g_{виб}k} - \bar{y}_{g_{виб}}). \qquad (3.18)$$

Формули (3.16) – (3.18) справедливі при умові, якщо у всіх рядках ВОМ опинилось однакове число паралельних дослідів, тобто $m_{gвиб} = m_{виб} = \text{const}$.

Критичне значення t-критерію вибирають з додатка В при вибраному рівні значимості q для числа ступенів свободи

$$\nu_{зн} = \nu_{вос} = N_{виб}\, m(m_{виб} - 1). \qquad (3.19)$$

Якщо $a_{iвиб} > a_{iвиб.кр}$, оцінки $a_{iвиб}$ вважають статистично значимими.

Якщо число паралельних дослідів в $g_{виб}$-х рядках ВОМ стали неоднаковими, то замість (3.17) потрібно використати формулу

$$s^2_{відт.виб}\{y\} = \frac{1}{\sum\limits_{g_{виб}=1}^{N} m_{g_{виб}}} \sum_{g_{виб}=1}^{N_{виб}} m_{g_{виб}} s^2_{g_{виб}}\{y\}, \qquad (3.20)$$

а замість (3.16) – формулу

$$s^2\{a_{i_{виб}}\} = \frac{1}{m \sum\limits_{g_{виб}=1}^{N_{виб}} m_{g_{виб}}} s^2_{відт.виб}\{y\}, \qquad (3.21)$$

причому число ступенів свободи для $t_{кр}$ складає

$$\nu_{зн} = \nu_{вос} = m \sum_{g_{виб}=1}^{N_{виб}} m_{g_{виб}} - N_{виб}. \qquad (3.22)$$

В обох випадках губиться $N_{виб}$ ступенів свободи на облік $N_{виб}$ порядкових дисперсій $s^2_{gвиб}\{y\}$ ($g = 1, 2, ..., N_{виб}$) у ВОМ. Якщо оцінки коефіцієнтів $a_{iвиб}$ та $a_{irвиб}$, обчислені за допомогою ВОМ, виявились значимими, то коректування виконують (якщо фактори z_i, що виділяються, та фактор взаємодії $z_i z_r$ вирішено стабілізувати на рівнях $z_i = -1$) за формулою

$$y_g^{\,I} = y_g - 2\sum_{i=1}^{n_{виб}} a_{i_{виб}} - 2\sum_{\substack{i,r=1 \\ i<r}}^{n_{виб}} a_{ir_{виб}}, \qquad (3.23)$$

де $n_{виб}$ – число факторів, для яких побудована ВОМ; i, r – номери факторів, включених у ВОМ. Формула (3.8) є окремим випадком (3.23).

Перевірку адекватності отриманої неповної квадратичної моделі (3.7) виконують за допомогою F-критерію Фішера, для чого необхідно отримати оцінку дисперсії відтворюванності за вихідною МП. Методика перевірки адекватності викладена в темі 2.

ЗАВДАННЯ І ПОРЯДОК ВИКОНАННЯ РОБОТИ

1. Задати модель об'єкта №6 з 8 входами x_1, x_2, ..., x_8, середнє квадратичне відхилення помилки відтворення дослідів $\sigma_y = 2$.
2. Скласти матрицю планування з 16 рядків для 8 лінійних факторів.
3. Провести експеримент відсіювання у районі базової точки: x_1=5,0; x_2=6,0; x_3=5,0; x_4=4,0; x_5=7,0; x_6=5,0; x_7=5,0; x_8=3,0. Прийняти інтервал варіювання Δx_i=0,3(i= 1, ... ,8).
4. За допомогою діаграми розсіювання, а також за допомогою вибіркових ортогональних матриць послідовно виділити найбільш істотні фактори.
5. Вибрати найбільш істотні парні взаємодії за виглядом діаграм розсіювання основних факторів, скласти для них матрицю планування у вільних стовпцях стандартного бланка, побудувати діаграму розсіювання.
6. Розрахувати оцінки коефіцієнтів a_i, a_{ir} для найбільш істотних факторів і парних взаємодій і скласти неповну квадратичну модель об'єкта.
7. Провести статистичне оцінювання результатів.

ФОРМУЛИ ДЛЯ РОЗРАХУНКІВ

1. $B_{z_i} = Me\{y\}\big|_{z_i=+1} - Me\{y\}\big|_{z_i=-1}$. 2. $y_g^1 = y_g - B_{z_i} - B_{z_r}$.

3. $a_{i_{виб}} = \dfrac{1}{N_{виб}} \sum\limits_{g_{виб}=1}^{N_{виб}} z_{ig_{виб}} y_{g_{виб}}$ $(i = 1, 2, ..., k)$.

4. $y_g^1 = y_g - 2a_i - 2a_r$. 5. $a_i = 0{,}5B_{z_i}$. 6. $a_{ir} = 0{,}5B_{z_i z_r}$.

7. $a_0 = \dfrac{1}{N_{виб}} \sum\limits_{g_{виб}=1}^{N_{виб}} \bar{y}_{g_{виб}}$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. У чому полягають основні задачі МВБ?
2. У чому полягає сутність МВБ?
3. Які основні теоретичні передумови застосування МВБ?
4. У чому полягають переваги і хиби МВБ у порівнянні з ПФЕ або ДФЕ?
5. Що таке насичене і перенасичене планування?

6. За яких умов вплив істотних змінних може бути незалежно оцінений кількісно?
7. Як будується матриця планування для експериментів відсіювання за МВБ?
8. Які заходи варто здійснити, щоб при виконанні МВБ одержати максимальну кількість незалежних оцінок коефіцієнтів рівняння регресії і провести їхній статистичний аналіз?
9. Як будується діаграма розсіювання за результатами експерименту?
10. Що таке внесок фактора і як він обчислюється?
11. Чому в МВБ за центр групування приймається медіана?
12. У чому полягає процедура послідовного виділення найбільш істотних факторів, які формули при цьому застосовуються?
13. Для чого застосовуються вибіркові ортогональні МП і як вони будуються?
14. Що таке "точки, що виділяються" і яке значення вони мають у МВБ?
15. Коли закінчується процес виділення найбільш істотних змінних?
16. Як за даними експерименту при МВБ оцінити неповну квадратичну модель об'єкта і провести її статистичний аналіз?
17. Як і для чого в МВБ використовується таблиця випадкових чисел?
18. Як за діаграмою розсіювання, побудованою для лінійних факторів, оцінити найбільш істотні парні взаємодії?
19. Яким повинно бути співвідношення внесків усіх факторів, щоб можна було успішно застосувати МВБ?
20. Чому коефіцієнт рівняння регресії в два рази менший від внеску?
21. Як при використанні МВБ оцінити дисперсію відтворюваності?
22. Як при використанні МВБ оцінити дисперсію коефіцієнтів рівняння регресії?

Тема 4 ДИСПЕРСІЙНИЙ АНАЛІЗ

Мета – *ознайомлення з призначенням і передумовами застосування методу дисперсійного аналізу, з його ідеєю і обчислювальним алгоритмом для опрацювання емпіричної інформації.*

КОРОТКІ ТЕОРЕТИЧНІ ВІДОМОСТІ

4.1 Постановка задачі

У багатьох областях практичної діяльності зустрічаються об'єкти дослідження, стан яких визначається вхідними змінними (факторами), що не мають кількісного опису. Такими факторами можуть бути некеровані і керовані змінні, що з якихось причин не дозволяють робити їхні вимірювання у даному експерименті, а також ті неконтрольовані змінні, рівні варіювання котрих можна довільно вибирати і фіксувати в часі. Для вивчення впливу факторів подібного роду на вихідну функцію об'єкта (відгук), їх загального оцінювання, ранжування і виділення серед них істотних, очевидно, непридатні методи відсіювання керованих кількісних факторів і метод регресійного аналізу некерованих факторів, оскільки ці методи передбачають вимірювання рівнів факторів, що досліджуються.

Приклади. 1. Розглянемо процес вимірювання групою операторів будь-якої фізичної величини декількома приладами (декількома методами), причому кожний оператор за допомогою кожного приладу проводить деяке число паралельних вимірювань. Усереднення спостережень, що відносяться до кожного з усіх можливих поєднань прилад–оператор, дає множину середніх арифметичних. Їх розсіювання може бути пов'язане з випадковою похибкою, систематичною похибкою приладу (методу) і впливом оператора. Потрібно визначити, наскільки істотно впливають на результат вимірювання два фактори: прилад (метод) та дії оператора.

2. Аналогічна задача виникає при оброблянні деталі паралельно на декількох станках автоматичної лінії. Вимагається встановити, чи однакові середні розміри деталей, що виготовляються на різних станках, таким чином оцінити, чи суттєво впливає фактор індивідуальності станка на процес обробляння.

3. При використанні деталей з декількох партій треба визначити, чи суттєво різняться параметри деталей різних партій.

4. При нестабільності величини вихідного показника в часі необхідно оцінити, наскільки суттєвий вплив повільного часового дрейфу на фоні випадкових похибок спостережень.

5. При налагоджуванні декількома операторами радіо-, телеі радіолокаційної апаратури візуальними і акустичними метода-

ми треба оцінити вплив на результат налагоджування трьох факторів: апаратури, методу і дій оператора.

Розглянемо тепер постановку задачі в загальному вигляді.

Дано: 1) відгук y може залежати (з фізичних причин) від n незалежно керованих факторів x_1, x_2, ... x_i, ... , x_n, що не мають кількісного опису, та їх парних взаємодій; 2) кожний фактор x_i може варіюватися на u_i рівнях; 3) повний факторний експеримент складається з $U= u1, u2, ..., u_i, ...u_n$ серій незалежних спостережень за числом всіх можливих неповторних сполучень рівнів n факторів; 4) кожна j-та серія містить m_j спостережень $y_{j1}, y_{j2}, ..., y_{ji}, ..., y_{jmj}$ паралельних дослідів.

Потрібно: визначити якою мірою істотний, на фоні випадкових похибок, вплив того чи іншого фактора x_i чи взаємодії факторів на відгук y; провести порівняння з іншими факторами та виділити найбільш важливі.

Припущення, на яких базується дисперсійний аналіз: спостереження відгуку y – нормально розподілена випадкова величина з центром розподілу $M\{y\}=\varphi(x_1, x_2, ..., x_i, ..., x_n)$. Таким чином, фактори визначають величину y лише в середньому, залишаючи простір для випадкових похибок спостережень, що підкоряються нормальному розподілу; дисперсія одиничного спостереження σ_ε^2, обумовлена випадковими похибками ε, постійна в усіх дослідах і не залежить від x_1, x_2, ..., x_i, ..., x_n. Таким чином дисперсії $\sigma_j^2\{y\}$ дорівнюють (j = 1, 2, ... , N), а їх вибіркові оцінки $\sigma_j^2\{y\}$ однорідні – умова відтворюваності дослідів з рівною точністю. Кожне з цих припущень необхідно перевірити за результатами спостережень за експериментом, що аналізується.

Із даних задачі та вказаних припущень бачимо, що чим більший вплив деякого фактора x_i на відгук y, тим більше розходження між середніми арифметичними відгуку y_j (j=1, 2, ..., u_i) в серіях паралельних спостережень, проведених при різних рівнях варіювання фактора x_i. Статистична значимість такого розходження вказує на суттєвий вплив фактора. При двох серіях спостережень порівняння середніх і перевірка нуль-гіпотези про значимості їх різниці проводяться за допомогою t-критерію Стьюдента. В сформульованій задачі вимагається одночасно порівняти довільно велике число середніх і на основі цього зробити висновок про суттєвість впливу того чи іншого фактора.

4.2 Ідея дисперсійного аналізу

Щоб мати можливість оцінювати вплив кожного фактора на відгук і порівнювати вплив різних факторів, встановимо деякий кількісний показник цього впливу. Нехай за відсутності помилок досліду ($\sigma_\varepsilon^2 = 0$) при варіюванні фактора x на u різних рівнях отримані достовірні значення $y_1, y_2, ..., y_j, ..., y_u$ відгуку y. Тоді в якості показника впливу фактора x приймаємо розмір, за аналогією зі звичайною дисперсією, що називається дисперсією фактора x, тобто

$$\sigma_x^2 = \frac{1}{u}\sum_{j=1}^{u}(y_j - \overline{y})^2, \qquad (4.1)$$

де $\overline{y} = \frac{1}{u}\sum_{j=1}^{u} y_j$. При цьому будемо мати на увазі, що числа y_j не є випадковими і тому дисперсія $\sigma_x{}^2$ не пов'язана ні з яким випадковим розміром, тому що ми приймаємо $\sigma_\varepsilon{}^2=0$. Вивчати вплив факторів за розмірами їхніх дисперсій зручно, оскільки це найпростіша міра розсіювання і до того ж аналогічна мірі впливу фактора випадкових розмірів, тобто дисперсії одиничного спостереження (відтворення) $\sigma_\varepsilon{}^2$. Завдяки цьому є можливість порівнювати вплив будь-якого досліджуваного фактора і фактора випадку. Таке дослідження факторів за їхніми дисперсіями називається дисперсійним аналізом.

Розглянемо ідею дисперсійного аналізу (ДА) на прикладі вивчення впливу одного фактора x на фоні випадкових похибок, коли дисперсія відтворюваності σ_ε^2 відома. При варіюванні фактора х на u рівнях в результаті спостережень отримаємо значення відгуку $y_1, y_2, ..., y_j, ..., y_u$, розсіювання яких можна характеризувати вибірковою дисперсією

$$S_0^2\{y\} = \frac{1}{u-1}\sum_{j=1}^{u}(y_j - \overline{y})^2, \qquad (4.2)$$

з числом ступенів свободи $\gamma_0 = u - 1$. Якщо відмінність $s_0^2\{y\}$ від σ_ε^2 незначна, то розкид спостережень, який вона характеризує, пов'язаний тільки із випадковими причинами і вплив фактора x несуттєвий. Якщо ж відмінність $s_0^2\{y\}$ від σ_ε^2 значна, то підвищений розкид спостережень викликається не тільки випадковими причинами, але ще й впливом фактора x, який тепер потрібно признати суттєвим. Оскільки в останньому випадку додаються впливи двох незалежних факторів: 1) випадкових причин (з дисперсією відгуку $\sigma_\varepsilon^2\{y\}$); 2) фактора x (з дисперсією σ_x^2), що призводить до загального розсіювання спостережень, – то загальна дисперсія є сумою $\left(\sigma_\varepsilon^2 + \sigma_x^2\right)$ двох вказаних, а її оцінка

$$S_0^2 = \overline{S}_0^2\{y\} \approx \sigma_\varepsilon^2 + \sigma_x^2, \tag{4.3}$$

звідки дисперсія фактора визначається виразом

$$\sigma_x^2 \approx s_0^2 + \sigma_\varepsilon^2. \tag{4.4}$$

У загальному випадку, коли дисперсія відтворюваності σ_ε^2 відома, схема ДА повинна дозволити знайти її оцінку наряду з оцінками дисперсій факторів, що вивчаються. З цією метою планується проведення серій паралельних дослідів при кожному із всіх можливих сполучень рівнів факторів. Таким чином, основна ідея ДА полягає в розкладанні оцінки загального розсіювання відгуку на складові, що залежать: від випадкових причин; від кожного з факторів; від їх взаємодії, а також від оцінювання статистичної значимості дисперсій останніх з урахуванням похибки відтворюваності досліду.

В цій темі ми розглянемо лише найпростіші способи застосування ДА, техніка проведення якого доволі різноманітна.

4.3 Однофакторний дисперсійний аналіз

Нехай вивчається вплив тільки одного фактора x. Представимо в табл. 4.1. результати експерименту із $u \times m$ спостережень y_{jl}.

Таблиця 4.1.

j \ l	1	2	…	l	…	m	$\overline{y}_j = \frac{1}{m}\sum_{l=1}^{m} y_{jl}$
1	y_{11}	y_{12}	…	y_{1l}	…	y_{1m}	y_1
2	y_{21}	y_{22}	…	y_{2l}	…	y_{2m}	y_2
…	…	…	…	…	…	…	…
j	y_{j1}	y_{j2}	…	y_{jl}	…	y_{jm}	y_j
…	…	…	…	…	…	…	…
u	y_{u1}	y_{u2}	…	y_{ul}	…	y_{um}	y_u
							$y = \frac{1}{um}\sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl} = \frac{1}{u}\sum_{j=1}^{u}\overline{y}_j$

де j – порядковий номер рівня варіювання фактора x (j=1,2,…,u); l – порядковий номер паралельного досліду в серії на кожному j – рівні (l=1,2,…,m_j). Для спрощення викладок, не порушуючи загальності висновків, розглянемо з початку випадок рівночислових серій спостережень на всіх рівнях, тобто $m_j = m = const$. При розташуванні спостережень в табл. 4.1 їх розсіювання між стовпцями обумовлюється похибкою відтворюваності, а розсіювання між рядками – додатковою дією фактора, що вивчається. Обчислимо середнє арифметичне $\overline{y}_j$ серій із m повторних спостережень для кожного j-го рівня фактора

$$\overline{y}_j = \frac{1}{m}\sum_{l=1}^{m} y_{jl} \tag{4.5}$$

і загальне середнє арифметичне $\overline{\overline{y}}$ всіх $u \times m$ спостережень по всіх u рівнях

$$\overline{\overline{y}} = \frac{1}{um}\sum_{j=1}^{m}\sum_{l=1}^{m} y_{jl} = \frac{1}{u}\sum_{j=1}^{m} y_j \,. \tag{4.6}$$

Розсіювання окремих спостережень відносно загального середнього обумовлено дією як випадкових факторів, так і впливом фактора x. Дія фактора випадковості проявляється в розсіюванні (з дисперсією σ_ε^2) спостережень серій паралельних дослідів y_{jl} на кожному рівні x_j навколо середнього арифметичного $\overline{y}_j$ своєї серії. Вплив фактора x (з дисперсією σ_x^2) викликає підвищене розсіювання середніх арифметичних $\overline{y}_j$ серій відносно загального середнього $\overline{\overline{y}}$. Кожне з цих трьох розсіювань можна охарактеризувати відповідною сумою квадратів відхилень.

1. Суми квадратів відхилень. У відповідності з основною ідеєю ДА розкладемо загальну суму квадратів відхилень спостережень y_{jl} від загального середнього $\overline{\overline{y}}$ на дві складові суми, одна з яких характеризує вплив фактора випадковості, а друга – фактора мінливості x

$$\begin{aligned}
S_0 &= \sum_{j=1}^{u}\sum_{l=1}^{m}(y_{jl}-\overline{\overline{y}})^2 = \sum_{j=1}^{u}\sum_{l=1}^{m}(y_{jl}-\overline{y}_j+\overline{y}_j-\overline{\overline{y}})^2 = \\
&= \sum_{j=1}^{u}\sum_{l=1}^{m}(y_{jl}-\overline{y}_j)^2 + 2\cdot\sum_{j=1}^{u}\sum_{l=1}^{m}(y_{jl}-\overline{y})\cdot(\overline{y}_j-\overline{\overline{y}}) + \sum_{j=1}^{u}\sum_{l=1}^{m}(\overline{y}_j-\overline{\overline{y}})^2 = \\
&= \sum_{j=1}^{u}\sum_{l=1}^{m}(y_{jl}-\overline{y}_j)^2 + m\cdot\sum_{j=1}^{m}(\overline{y}_j-\overline{\overline{y}})^2 = S_\varepsilon + S_x,
\end{aligned} \tag{4.7}$$

внаслідок того, що

$$\sum_{j=1}^{m}\sum_{l=1}^{m}(y_{jl}-\overline{y})\cdot(\overline{y}_j-\overline{\overline{y}}) = \sum_{j=1}^{u}(\overline{y}_j-\overline{\overline{y}})\sum_{l=1}^{m}(y_{jl}-\overline{y}_j) = 0\,, \tag{4.8}$$

оскільки $\sum_{j=1}^{m}(y_{jl}-\overline{y}_j) = 0$, як сума відхилень l-х спостережень j-ої серії від середнього тієї ж серії. Суми S_0, S_ε, S_x, що входять у співвідношення (4.7), означають:

$$S_0 = \sum_{j=1}^{u} \sum_{l=1}^{m} (\bar{y}_j - \bar{\bar{y}})^2 \tag{4.9}$$

S_0 – «загальна» сума квадратів відхилень окремих спостережень y_{jl} від загального середнього $\bar{\bar{y}}$. Вона характеризує розсіювання спостережень в результаті дії обох факторів, як випадковості ε, так і того, що вивчається, x;

$$S_\varepsilon = \sum_{j=1}^{u} \sum_{l=1}^{m} (\bar{y}_{jl} - \bar{y}_j)^2 \tag{4.10}$$

S_ε– сума квадратів відхилень «всередині серій», тобто сума квадратів різниці між окремими спостереженнями y_{jl} і середнім значенням $\bar{y}_j$ відповідної j-тої серії. Вона характеризує залишкове розсіювання випадкових похибок дослідів, тобто їхню відтворюваність;

$$S_x = \sum_{j=1}^{m} (\bar{y}_j - \bar{\bar{y}})^2 \tag{4.11}$$

S_x– сума квадратів відхилень «між серіями» чи розсіювання по рівнях фактора x, тобто зважена з урахуванням числа m паралельних спостережень в кожній серії сума квадратів різниць між середніми $\bar{y}_j$ окремих серій і загальним середнім $\bar{\bar{y}}$ по всіх сукупностях спостережень. Сума S_x/m характеризує розсіювання середніх $\bar{y}_j$ серій за рахунок випадкових причин (з дисперсією σ_ε^2/m для середніх серій) та фактора, що досліджується (з дисперсією σ_ε^2).

2. **Оцінки дисперсій**. Припустимо, що вплив фактора x на відгук відсутній, тобто нуль-гіпотеза про однорідність $\bar{\bar{y}}_j (j=1,2, \ldots, u)$ вірна. Тоді всі u серії паралельних спостережень можна розглядати як випадкові вибірки із одної і тої ж нормальної генеральної сукупності і, таким чином:

1) незміщена загальна оцінка дисперсій відтворюваності σ_ε^2 по всіх um спостереженнях визначається з виразу

$$S_0^2\{y\} = \frac{1}{um-1} \sum_{j=1}^{u} \sum_{l=1}^{m} (y_{jl} - \bar{y})^2 = \frac{S_0}{um-1} \approx \sigma_\varepsilon^2 \tag{4.12}$$

з кількістю ступенів свободи $\nu_0 = um - 1$;

2) вибіркова дисперсія розсіювання «всередині серій», чи залишкова оцінка дисперсії відтворюваності σ_ε^2, знаходиться як середнє з вибіркових дисперсій по кожній серії окремо

$$S_\varepsilon^2\{y\} = \frac{1}{u}\sum_{j=1}^{u}\frac{1}{m-1}\sum_{l=1}^{m}(y_{jl} - \overline{\overline{y}})^2 = \frac{S_\varepsilon}{u(m-1)} \approx \sigma_\varepsilon^2 \tag{4.13}$$

з кількістю ступенів свободи $\nu_\varepsilon = u(m-1)$;

3) вибіркова дисперсія середніх по серіях служить незміщеною оцінкою дисперсії σ_ε^2 / m, з якою нормально розподілені незалежні один від одного середні j-х серій

$$S_x^2\{\overline{y}_j\} = \frac{1}{u-1}\sum_{j=1}^{u}(\overline{y}_j - \overline{\overline{y}})^2 = \frac{S_x}{m(u-1)} \approx \frac{\sigma_\varepsilon^2}{m} \tag{4.14}$$

з кількістю ступенів свободи $\nu_x = u - 1$. Звідки легко отримаємо третю оцінку дисперсії відтворюваності, вибіркову дисперсію розсіювання «між серіями»

$$S_x^2\{y\} = \frac{S_x}{u-1} \approx \sigma_\varepsilon^2 \tag{4.15}$$

Підрахунок чисел ступенів свободи перевіряється за допомогою співвідношення $\nu_0 = \nu_\varepsilon = \nu_x$;

4) в результаті більш глибокого аналізу можна показати, що S_ε і S_x незалежні один від одного.

Із сказаного видно, що при відсутності впливу фактора x оцінки S_0^2, S_ε^2, S_x^2 однорідні, оскільки є оцінками однієї і тієї ж генеральної дисперсії σ_ε^2.

Припустимо тепер, що вплив фактора x на відгук суттєвий, тобто нуль-гіпотеза про однорідність $\overline{y}_j$ ($j = 1, 2, \ldots, u$) невірна. Тоді u серії паралельних спостережень можна розглядати як випадкові вибірки u незалежних нормально розподілених випадкових величин з однією і тією ж дисперсією σ_ε^2 і різними генеральними центрами розподілення $C_1, C_1, \ldots, C_j, \ldots, C_n$ і, отже:

1) вибіркова дисперсія S_0^2 характеризує вплив як фактора випадковості ε, так і фактора x, тобто

$$S_0^2 \approx \sigma_\varepsilon^2 + \sigma_x^2; \tag{4.16}$$

2) оскільки сума S_ε не змінюється при зміні y_{jl} на y_{jl}-C_j, де C_j – генеральний центр розподілення цільової функції y при стабілізації фактора x на j-у рівні, то вибіркова дисперсія S_ε^2 також не змінюється і є незміщеною оцінкою для генеральної дисперсії відтворюваності σ_ε^2, тобто

$$S_\varepsilon^2 \approx \sigma_\varepsilon^2; \tag{4.17}$$

3) оскільки сума S_x враховує не тільки випадкові, але й систематичні розходження між середніми серій та збільшується за

рахунок впливу фактора x, дисперсія $S^2\{\overline{y}_j\}$ при цьому також збільшується та перестає служити оцінкою тільки для σ_ε^2/m, тобто

$$S^2\{\overline{y}_j\} \approx \frac{\sigma_\varepsilon^2}{m} + \sigma_x^2,$$

звідки випливає, що

$$S_x^2\{y\} \approx \sigma_\varepsilon^2 + m\sigma_x^2; \tag{4.18}$$

незалежність S_ε та S_x один від одного зберігається.

Таким чином, для дисперсії фактора x тепер можна дати дві оцінки:

$$\sigma_x^2 \approx \sigma_0^2 - \sigma_\varepsilon^2 \approx S_0^2 - S_\varepsilon^2, \tag{4.19}$$

$$\sigma_x^2 \approx \frac{1}{m}(S_x^2 - \sigma_\varepsilon^2) \approx \frac{1}{m}(S_x^2 - S_\varepsilon^2). \tag{4.20}$$

Перша оцінка менш точна через похибки величин S_0^2 та S_ε^2. Точність другої вища, оскільки вибіркові дисперсії входять в неї поділеними на m. Із зробленого другого припущення видно, що при впливі фактора x оцінки $S_0^2, S_\varepsilon^2, S_x^2$ неоднорідні. Отже, після порівняння цих вибіркових дисперсій, можна прийняти рішення про справедливість першого чи другого припущення відносно істотності впливу фактора x (з дисперсією σ_x^2) на відгук. Враховуючи точність виразів (4.19), (4.20), для перевірки нуль-гіпотези H_0 : $\sigma_x^2 = 0$ будемо порівнювати вибіркові дисперсії S_x^2, S_ε^2.

3. **Оцінювання впливу фактора**. Для того, щоб вплив фактора x був признаний істотним ($\sigma_x^2 > 0$), необхідно і достатньо, щоб оцінка дисперсії S_x^2 значно відрізнялась від S_ε. Перевірку нуль-гіпотези про однорідність цих вибіркових дисперсій можна втілити за допомогою критерію Фішера

$$F = S_x^2/S_x^2. \tag{4.21}$$

Якщо обчислене за результатами спостережень дисперсійне відношення F перевищує табличне $F_q(\nu_x,\nu_\varepsilon)$, що знайдене за розподіленням Фішера для обраного рівня значимості q при відповідних ступенях свободи ν_x та ν_ε , то вплив фактора x признається істотним ($\sigma^2{}_x > 0$), і, навпаки, якщо $F \le F_q(\nu_x,\nu_\varepsilon)$, то – несуттєвим ($\sigma^2{}_x = 0$). В дисперсійному аналізі перевіряють нуль-гіпотезу H_0: $M\{S_x^2\} = M(S_\varepsilon^2) = \sigma_\varepsilon^2$ при альтернативі H_1: $M\{S_x^2\} > \sigma_\varepsilon^2$, тому користуються однобічним F-критерієм.

Слід мати на увазі, що дисперсійний аналіз спостережень експерименту дозволяє оцінювати вплив фактора лише в цілому і, що висновки, які отримані за його допомогою, відносяться тільки до даного експериментального матеріалу при даній його систематизації. Так, наприклад, при зміні діапазону варіювання фактора, що вивчається, чи основної (базової) точки, оцінка впливу останнього може змінитись.

Якщо вплив фактора x вважається неістотним, то дисперсію відтворюваності σ_ε^2 можна оцінити вибірковою загальною дисперсією $S^2{}_0$, яка має на $(u - 1)$ ступенів свободи більше ніж S_ε^2. Якщо ж вплив фактора x вважається істотним, то за результатами спостережень можна оцінити:

1) дисперсію вибіркової відтворюваності σ_ε^2 – залишковою дисперсією

$$S_\varepsilon^2 = \frac{S_\varepsilon}{u(m-1)} \approx \sigma_\varepsilon^2 \quad \text{тобто} \quad M\{S_\varepsilon^2\} = \sigma_\varepsilon^2 \tag{4.22}$$

і визначити довірчий інтервал для σ_ε^2 по χ^2-розподіленню з $u\times(m - 1)$ ступенями свободи;

2) дисперсію фактора x за формулою

$$\sigma_x^2 \approx \frac{1}{m}(S_x^2 - \sigma_\varepsilon^2); \tag{4.23}$$

3) розходження δ_C^2 центрів серій, обумовлене впливом фактора x. Оскільки

$$S_x^2 = \frac{S_x}{u-1}, \tag{4.24}$$

то можна показати, що

$$M\{S_x^2\} = \sigma_\varepsilon^2 + \frac{m}{u-1}\sum_{j=1}^{u}(C_j - \overline{C})^2, \tag{4.26}$$

де $\overline{C} = \frac{1}{u}\sum_{j=1}^{u} C_j$ – середнє значення із центрів розподілення C_j, чи

$$M\left\{\frac{u-1}{um}\left(S_x^2 - S_\varepsilon^2\right)\right\} = \frac{1}{u}\sum_{j=1}^{u}(C_j - \overline{C})^2 = \delta_C^2. \tag{4.27}$$

Оцінкою величини δ_C^2 служить вибіркова характеристика

$$d_x^2 \approx \frac{u-1}{um}(S_x^2 - \sigma_\varepsilon^2);$$

4) розходження $C_j - C_g$ між центрами будь-яких двох серій. Оскільки статистика

$$t = \sqrt{\frac{m}{2}}\,\frac{\left(\bar{y}_j - \bar{y}_g\right) - \left(C_j - C_g\right)}{S_\varepsilon} \tag{4.28}$$

відповідає розподіленню Стьюдента з кількістю ступенів свободи $\nu_\varepsilon = u(m-1)$, то інтервал

$$\left(y_j - y_g - t_{q:u(m-1)}\frac{S_\varepsilon}{\sqrt{m/2}};\; y_j - y_g + t_{q:u(m-1)}\frac{S_\varepsilon}{\sqrt{m/2}}\right)$$

служить надійним $(1-q)$ 100% інтервалом для $C_j - C_g$.

4. Випадок нерівночисленних серій спостережень. Вище ми розглянули випадок рівночисленних серій спостережень на всіх рівнях фактора x. Ця обставина несуттєва для теорії дисперсійного аналізу, і тому при різному числі паралельних спостережень на різних j-рівнях схема приведення та основні прийоми аналізу залишаються попередніми. Змінюється лише вигляд виразів:

1). загальне число спостережень

$$M = \sum_{j=1}^{a} m_j; \tag{4.29}$$

2). суми спостережень по серіях

$$Y_l = \sum_{l=1}^{m_j} y_{lt}; \tag{4.30}$$

3). середні в серіях

$$\bar{y}_j = \frac{1}{m_j}\sum_{l=1}^{m_j} y_{jt} = \frac{Y_l}{m_j}; \tag{4.31}$$

4). загальне середнє

$$\bar{y} = \frac{1}{M}\sum_{j=1}^{a}\sum_{l=1}^{m_j} = \frac{1}{M}\sum_{j=1}^{a} Y_j; \tag{4.32}$$

5). співвідношення для сум

$$S_0 = \sum_{j=1}^{a}\sum_{l=1}^{m_j}\left(y_{jl} - \bar{y}\right)^2 = \sum_{j=1}^{a}\sum_{l=1}^{m_j}\left(y_{jt} - \bar{y}_j\right)^2 +$$

$$+ \sum_{l=1}^{a} m_j\left(\bar{y}_j - \bar{\bar{y}}\right)^2 = S_\varepsilon + S_x; \tag{4.33}$$

6). співвідношення для числа ступенів свободи

$$\nu_0 = M - 1 = \left(M - u\right) + \left(u - 1\right) = \nu_\varepsilon - \nu_x; \tag{4.34}$$

7). дисперсія фактора x при суттєвому впливі факторів обчислюється за формулою

$$\sigma_x^2 \approx \frac{M(u-1)}{M^2 - \sum\limits_{l=1}^{u} m_l^2}\left(S_x^2 - S_\varepsilon^2\right). \tag{4.35}$$

5. Розрахункові формули для сум. Алгоритм однофакторного дисперсійного аналізу спрощується, якщо для розрахунку сум квадратів відхилень використовувати перетворення

$$\sum_{r=1}^{M}\left(y_r - \bar{y}\right)^2 = \sum_{r=1}^{M} y_r^2 - \frac{1}{M}\left(\sum_{r=1}^{M} y_r\right)^2. \tag{4.36}$$

Тоді для сум отримуються розрахункові формули:

ё

$$= \sum_{j=1}^{u}\sum_{l=1}^{m_j} y_{jl}^2 - \frac{1}{M}\left(\sum_{j=1}^{u} Y_j\right)^2 = Q_1 - Q_3; \tag{4.37}$$

$$S_\varepsilon = \sum_{j=1}^{u}\sum_{l=1}^{m_j}\left(y_{jl} - \bar{y}_j\right)^2 = \sum_{j=1}^{u}\sum_{l=1}^{m_j} y_{jl}^2 - \sum_{j=1}^{u}\frac{1}{m_j}\left(\sum_{l=1}^{m_j} y_{jl}\right)^2 =$$

$$= \sum_{j=1}^{u}\sum_{i=1}^{m_j} y_{ji}^2 - \sum_{j=1}^{u}\frac{1}{m_j} Y_j^2 = Q_1 - Q_2; \tag{4.38}$$

$$S_x = \sum_{j=1}^{u}\sum_{l=1}^{m_j}\left(\bar{y}_j - \bar{\bar{y}}\right)^2 = \sum_{j=1}^{u}\frac{1}{m_j}\left(\sum_{l=1}^{m_j} y_{jl}\right)^2 - \frac{1}{M}\left(\sum_{j=1}^{u}\sum_{l=1}^{m_j} y_{jl}\right)^2 =$$

$$= \sum_{l=1}^{u}\frac{1}{m_j} Y_l^2 - \frac{1}{M}\left(\sum_{l=1}^{u} Y_j\right)^2 = Q_2 - Q_3. \tag{4.39}$$

Таким чином, для проведення однофакторного ДА достатньо попередньо обчислити:

1) суми спостережень по серіях

$$Y_j = \sum_{l=1}^{m_j} y_{jl}; \tag{4.40}$$

2) суму квадратів всіх M спостережень

$$Q_1 = \sum_{j=1}^{u}\sum_{l=1}^{m_j} y_{jl}^2; \tag{4.41}$$

3) суму квадратів висновків по серіях, поділених на число спостережень в серії,

$$Q_2 = \sum_{j=1}^{u}\frac{1}{m_j} Y_l^2; \tag{4.42}$$

4) квадрат загального висновку, поділений на число всіх спостережень,

$$Q_3 = \frac{1}{M}\left(\sum_{j=1}^{u} Y_j\right)^2. \quad (4.43)$$

6. Приклад. Розглянемо приклад використання однофакторного ДА при спостереженні залежності довговічності y (год) електричних ламп від технології виготовлення (фактор x). Припустимо, що виконуються припущення ДА, тобто випадкова величина довговічності має нормальний розподіл і вплив технології виготовлення електричних ламп не впливає на дисперсію σ_ε^2 величини y, але може викликати розбіжність середніх значень. Були відібрані нерівночислові серії зразків з чотирьох партій продукції (u=4). Результати спостережень приведені в табл. 4.2, при чому для спрощення обчислень всі дані зменшені на одну й ту ж величину (1500 год), оскільки на значеннях дисперсій це не відбивається. Підрахунки, виконані в цій самій таблиці, зрозумілі без пояснень і служать для знаходження сум квадратів відхилень та вибіркових дисперсій. Ці величини наведені в табл. 4.3 однофакторного ДА.

Таблиця 4.2

l / j	1		2		3		4		5	
	y_{j1}	y^2_{j1}	y_{j2}	y^2_{j2}	y_{j3}	y^2_{j3}	y_{j4}	y^2_{j4}	y_{j5}	y^2_{j5}
1	100	10000	110	12100	150	22500	180	32400	200	40000
2	80	6400	140	19600	140	19600	200	40000	250	62500
3	-40	1600	50	2500	100	10000	120	14400	140	19600
4	10	100	20	400	30	900	70	4900	100	10000
$\sum_{j=1}^{u=4}$	150	18100	320	34600	420	53000	570	91700	690	132100

Продовження табл. 4.2

l / j	6		7		8		Y_j	$\frac{1}{m_j} Y_j^2$	$\sum_{i=1}^{m_j} y_{ji}^2$
	y_{j6}	y^2_{j6}	y_{j7}	y^2_{j7}	y_{j8}	y^2_{j8}			
1	200	40000	300	90000	150	22500	1240	219657	247000
2					140	19600	810	131220	148400
3	160	25600	240	57600	100	10000	1090	148500	233700
4	180	32400			30	900	410	28017	48700
$\sum_{j=1}^{u=4}$	540	98000	540	147600	420	53000	3550	Q_2= 527394	Q_1= 677500

Виконаємо перевірку сутності впливу фактора x, для чого знайдемо дисперсне відношення:

$$F = s_x^2 / s_\varepsilon^2 = 14227 / 6823 = 2{,}08.$$

Таблиця 4.3

Розсіювання	Сума квадратів відхилень	Число ступенів свободи	Вибіркова дисперсія	Компоненти генеральної дисперсії
Загальне	S_0=192788	$\nu_0=M-1=25$	$s^2{}_0$=7712	$\sigma^2{}_\varepsilon+\sigma^2{}_x$
Всередині серій	S_ε =150106	$\nu_\varepsilon=M-u=22$	$s^2{}_\varepsilon$=6823	$\sigma^2\varepsilon$
Між серіями	S_x=42682	$\nu_x=u-1=3$	$s^2{}_x$=14227	$\sigma^2{}_\varepsilon+m\sigma^2{}_x$

З додатку Г знаходимо, що значення F-критерію для 5%-вого рівня значимості q та чисел ступенів свободи $\nu_1=\nu_x=3$ та $\nu_2=\nu_\varepsilon=22$ складає $F_{0,05}(3;22)=3,05$. Оскільки $F<F_{0,05}(3;22)$, то вплив технології виготовлення на тривалість роботи електричних ламп в розглянутих чотирьох партіях може вважатися несуттєвим.

4.4 Двофакторний дисперсійний аналіз

Нехай вивчається вплив двох факторів x_1 і x_2, що діють одночасно. Представимо в таблиці 4.4 результати експерименту з $u_1 x u_2 x m$ спостережень y_{jgl}, де j-порядковий номер рівня варіювання фактора x_2 (g=1,2, ..., u_2); l – порядковий номер паралельного досліду в серії при кожному jg-му співвідношенні рівнів двох факторів (l=1,2, ..., m_{jg}). Для спрощення викладок розглянемо випадок рівночислових серій спостережень при всіх можливих співвідношеннях рівнів, тобто $m_{lg}=m=const$.

Обчислимо середні арифметичні $\bar{y}_{jg}$ серій з m повторних спостережень для кожного співвідношення j-го та g-го рівнів факторів x_1 та x_2

$$\bar{y}_{jg} = \frac{1}{m}\sum_{l=1}^{m} y_{jgl}; \tag{4.44}$$

середнє арифметичне $\bar{y}_j$ по рядках табл. 4.4 з $u_2 m$ спостережень для кожного j-го рівня фактора x_1

$$\bar{y}_j = \frac{1}{u_2 m}\sum_{g=1}^{u_2}\sum_{l=1}^{m} y_{jgl} = \frac{1}{u_2}\sum_{g=1}^{u_2}\bar{y}_{jg}; \tag{4.45}$$

середнє арифметичне $\bar{y}'_g$ по стовпцях табл. 4.4 з $u_1 m$ спостережень для кожного g-го рівня фактора x_2

$$\bar{y}'_g = \frac{1}{u_1 m}\sum_{j=1}^{u_1}\sum_{l=1}^{m} y_{jgl} = \frac{1}{u_1}\sum_{j=1}^{u_1}\bar{y}_{jg}; \tag{4.46}$$

загальне середнє арифметичне $\bar{y}$ всіх $M=u_1u_2m$ спостережень по всіх u_1u_2 співвідношеннях рівнів

$$\bar{\bar{y}} = \frac{1}{u_1u_2m}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m} y_{jgl} = \frac{1}{u_1u_2}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\bar{y}_{jg} =$$

$$= \frac{1}{u_1}\sum_{j=1}^{u_1}\bar{y}_j = \frac{1}{u_2}\sum_{g=1}^{u_2}\bar{y}'_g. \qquad (4.47)$$

При вказаному розташуванні спостережень їх розсіювання в кожній серії відносно середнього тієї ж серії обумовлено дією тільки випадкових причин (з дисперсією σ_ε^2). Розсіювання ж самих середніх в серіях по всіх можливих співвідношеннях рівнів x_1 і x_2 навколо загального середнього, окрім фактора випадковості, виливається впливом фактора взаємодії x_1x_2 (з дисперсією $\sigma_{x_1x_2}^2$). Окрім цих факторів на розсіювання середніх по рядках здійснює вплив тільки один фактор x_1 (з дисперсією $\sigma_{x_1}^2$), а на розсіювання середніх по стовпцях – тільки один фактор x_2 (з дисперсією $\sigma_{x_2}^2$), оскільки всі рівні іншого фактора в кожному з цих випадків усереднені.

1. Суми квадратів відхилень. У відповідності з основною ідеєю ДА розкладемо загальну суму S_0 квадратів відхилень спостережень від загального середнього на компоненти, що відповідають перерахованим вище факторам

$$S_0 = \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(y_{jgl} - \bar{\bar{y}}\right)^2 =$$

$$= \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(y_{jgl} - \bar{y}_{jg} + \bar{y}_{jg} - \bar{y}_j + \bar{y}_j - \bar{y}'_g + \bar{y}'_g - \bar{\bar{y}} + \bar{\bar{y}} - \bar{\bar{y}}\right)^2 =$$

$$= \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(y_{jgl} - \bar{y}_{jg}\right)^2 + \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(\bar{y}_j - \bar{\bar{y}}\right)^2 +$$

$$+ \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(\bar{y}'_g - \bar{\bar{y}}\right)^2 + \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(\bar{y}_{jg} - \bar{y}_j - \bar{y}'_g + \bar{\bar{y}}\right)^2 =$$

$$= S_\varepsilon + S_{x_1} + S_{x_2} + S_{x_1x_2}. \qquad (4.48)$$

S_0 – загальна сума квадратів, що характеризує розсіювання окремих спостережень y_{jgl} в загальній сукупності, за рахунок впливу всіх факторів;

$$S_\varepsilon = \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(y_{jgl} - \bar{y}_{jg}\right)^2 \qquad (4.49)$$

S_ε – сума квадратів відхилень «всередині серій», що характеризує розсіювання окремих спостережень y_{jgl} в серіях тільки за рахунок впливу фактора випадковості, оскільки на протягом серії фактори x_1 і x_2 залишаються незмінними;

Таблиця 4.4

Рівень g x_2 / Рівень j x_1	1	2	…	g	…	u_2	$\bar{y}_j = \frac{1}{u_2 m}\sum_{g=1}^{u_2}\sum_{l=1}^{m} y_{igl}$
1	y_{111} y_{112} … y_{11l} … y_{11m}	y_{121} y_{122} … y_{12l} … y_{12m}	… … … … … …	y_{1g1} y_{1g2} … y_{1gl} … y_{1gm}	… … … … … …	y_{1u21} y_{1u22} … y_{1u2l} … y_{1u2m}	y_1
2	y_{211} y_{212} … y_{21l} … y_{21m}	y_{221} y_{222} … y_{22l} … y_{22m}	… … … … … …	y_{2g1} y_{2g2} … y_{2gl} … y_{2gm}	… … … … … …	y_{2u21} y_{2u22} … y_{2u2l} … y_{2u2m}	y_2
…	…	…	…	…	…	…	…
j	y_{j11} y_{j12} … y_{j1l} … y_{j1m}	y_{j21} y_{j22} … y_{j2l} … y_{j2m}	… … … … … …	y_{jg1} y_{jg2} … y_{jgl} … y_{jgm}	… … … … … …	y_{ju21} y_{ju22} … y_{ju2l} … y_{ju2m}	y_j
…	…	…	…	…	…	…	…
u_1	y_{u111} y_{u112} … y_{u11l} … y_{u11m}	y_{u121} y_{u122} … y_{u12l} … y_{u12m}	… … … … … …	y_{u1g1} y_{u1g2} … y_{u1gl} … y_{u1gm}	… … … … … …	y_{u1u21} y_{u1u22} … y_{u1u2l} … y_{u1u2m}	y_{u1}
$\bar{y}'_j = \frac{1}{u_1 m}\sum_{j=1}^{u_1}\sum_{l=1}^{m} y_{igl}$	y'_1	y'_2	…	y'_g	…	y'_{u2}	$\bar{y}_j = \frac{1}{u_1 u_2 m}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m} y_{igl}$

$$S_{x_1} = \sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}\left(\bar{y}_j - \bar{\bar{y}}\right)^2 = u_2 m \sum_{j=1}^{u_1}\left(\bar{y}_j - \bar{\bar{y}}\right)^2 \qquad (4.50)$$

S_{x_1}– сума квадратів відхилень «між рядками». Сума $S_{x_1} / (u_2 m)$ характеризує розсіювання середніх $\bar{y}_j$ по рядках в результаті дії фактора випадковості (з дисперсією середнього для рядка $\sigma_\varepsilon^2 / (u_2 m)$), фактора x_1 (з дисперсією $\sigma_{x_1}^2$) і фактора взаємодії (з дисперсією середнього для рядка $\sigma_{x_1 x_2}^2 / u_2$);

$$S_{x_2} = \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \sum_{l=1}^{m} \left(\bar{y}'_g - \bar{\bar{y}} \right)^2 = u_1 m \sum_{g=1}^{u_2} \left(\bar{y}'_g - \bar{\bar{y}} \right)^2 \qquad (4.51)$$

S_{x_2} – сума квадратів відхилень «між стовпцями». Сума $S_{x_2} / (u_1 m)$ характеризує розсіювання середніх по стовпцях в результаті дії фактора випадковості (з дисперсією середнього для стовпця $\sigma_\varepsilon^2 / (u_1 m)$), фактора x_1 (з дисперсією $\sigma_{x_1}^2$) і фактора взаємодії (з дисперсією середнього для рядка $\sigma_{x_1 x_2}^2 / u_1$);

$$S_{x_1 x_2} = \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \sum_{l=1}^{m} \left(\bar{y}_{jg} - \bar{y}_j - \bar{y}'_g + \bar{\bar{y}} \right)^2 =$$

$$= m \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \left(\bar{y}_{jg} - \bar{y}_j - \bar{y}'_g + \bar{\bar{y}} \right)^2 \qquad (4.52)$$

$S_{x_1 x_2}$– сума квадратів відхилень «між серіями». Сума $S_{x_1 x_2 / m}$ характеризує розсіювання середніх $\bar{y}_{jg}$ серій в результаті дії фактора випадковості (з дисперсією середнього серії σ_ε^2 / m) і фактора взаємодії (з дисперсією $\sigma_{x_1 x_2}^2$).

2. Оцінки дисперсій. Кожна з сум квадратів S_0, S_ε, S_{x_1}, S_{x_2}, $S_{x_1 x_2}$, розділена на відповідне їй число ступенів свободи ν_0, ν_ε, ν_{x_1}, ν_{x_2}, $\nu_{x_1 x_2}$, дає незміщену оцінку відповідної дисперсії:

1) вибіркову загальну дисперсію по всіх $M = u_1 u_2 m$ спостереженням відгуку

$$s_0^2 \{y\} = \frac{1}{M-1} \sum_{j=1}^{u_1} \sum_{g=1}^{u_2} \sum_{l=1}^{m} \left(y_{jgl} - \bar{\bar{y}} \right)^2 =$$

$$= \frac{S_0}{M-1} \approx \sigma_\varepsilon^2 + \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_1 x_2}^2 \qquad (4.53)$$

з числом ступенів свободи $\nu_0 = u_1 u_2 m - 1 = M - 1$;

2) вибіркову дисперсію розсіювання «всередині серій» (остаточну оцінку), яка є середньозваженою оцінкою дисперсії по всіх $u_1 u_2$ серіях спостережень:

2) вибіркову дисперсію розсіювання «всередині серій» (остаточну оцінку), яка є середньозваженою оцінкою дисперсії по всіх $u_1 u_2$ серіях спостережень:

$$s_\varepsilon^2\{y\} = \frac{1}{u_1 u_2}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\frac{1}{m-1}\sum_{l=1}^{m}\left(y_{jgl} - \bar{y}_{jg}\right)^2 = \frac{S_\varepsilon}{u_1 u_2(m-1)} \approx \sigma_\varepsilon^2 \qquad (4.54)$$

з числом ступенів свободи $\nu_\varepsilon = u_1 u_2(m-1)$;

3) вибіркову дисперсію розсіювання «між рядками»

$$s_{x_1}^2\{y\} = \frac{u_2 m}{u_1 - 1}\sum_{j=1}^{u_1}\left(\bar{y}_j - \bar{\bar{y}}\right)^2 = \frac{S_{x_1}}{u_1 - 1} \approx \sigma_\varepsilon^2 + u_2 m\sigma_{x_1}^2 + m\sigma_{x_1x_2}^2 \qquad (4.55)$$

з числом ступенів свободи $\nu_{x_1} = u_1 - 1$;

4) вибіркову дисперсію розсіювання «між стовпцями»

$$s_{x_2}^2\{y\} = \frac{u_1 m}{u_2 - 1}\sum_{j=1}^{u_2}\left(\bar{y}'_g - \bar{\bar{y}}\right)^2 = \frac{S_{x_2}}{u_2 - 1} \approx \sigma_\varepsilon^2 + u_1 m\sigma_{x_2}^2 + m\sigma_{x_1x_2}^2 \qquad (4.56)$$

з числом ступенів свободи $\nu_{x_2} = u_2 - 1$;

5) вибіркову дисперсію розсіювання «між серіями»

$$s_{x_1x_2}\{y\} = \frac{m}{(u_1 - 1)(u_2 - 1)}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\left(\bar{y}_{jg} - \bar{y}_j - \bar{y}'_g + \bar{\bar{y}}\right)^2 =$$

$$= \frac{S_{x_1x_2}}{(u_1 - 1)(u_2 - 1)} \approx \sigma_\varepsilon^2 + m\sigma_{x_1x_2}^2 \qquad (4.57)$$

з числом ступенів свободи $\nu_{x_1x_2} = (u_1 - 1)(u_2 - 1)$.

Правильність підрахунку чисел ступенів свободи перевіряється за допомогою співвідношення

$$\nu_0 = \nu_\varepsilon + \nu_{x_1} + \nu_{x_2} + \nu_{x_1x_2}. \qquad (4.58)$$

3. Оцінка важливості впливу факторів та їх взаємодії. Аналіз важливості впливу факторів x_1, x_2 та їх взаємодію x_1x_2 виконують на підставі критерію Фішера при обраному рівні значимості q в такому порядку:

1. Вплив факторів x_1 і x_2 оцінюють відповідно до дисперсій:

$$\sigma_{x_1}^2 \approx \frac{1}{u_2 m}\left(s_{x_1}^2 - \sigma_\varepsilon^2 - m\sigma_{x_1x_2}^2\right) \approx \frac{1}{u_2 m}\left(s_{x_1}^2 - s_{x_1x_2}^2\right),$$

$$\sigma_{x_2}^2 \approx \frac{1}{u_1 m}\left(s_{x_2}^2 - \sigma_\varepsilon^2 - m\sigma_{x_1x_2}^2\right) \approx \frac{1}{u_1 m}\left(s_{x_2}^2 - s_{x_1x_2}^2\right) \qquad (4.59)$$

і визнають суттєвим ($\sigma_{x_1}^2 > 0; \sigma_{x_2}^2 > 0$), якщо відповідно буде значно відрізнятися $S_{x_1}^2$ від $S_{x_1x_2}^2$ та $S_{x_2}^2$ від $S_{x_1x_2}^2$, тобто, якщо відповідний емпіричний критерій більший критичного:

$$F_1 = s^2_{x_1} / s^2_{x_1x_2} > F_q\left(\nu_{x_1}; \nu_{x_1x_2}\right),$$
$$F_2 = s^2_{x_2} / s^2_{x_1x_2} > F_q\left(\nu_{x_2}; \nu_{x_1x_2}\right). \qquad (4.60)$$

В тому випадку, коли одне з цих дисперсійних відношень незначно відрізняється від одиниці, тобто вплив одного з факторів несуттєвий ($\sigma^2_{x_1} = 0$ або $\sigma^2_{x_2} = 0$), для дисперсії $\sigma^2_\varepsilon + m\sigma^2_{x_1x_2}$ отримуються дві однорідні оцінки відповідно $S^2_{x_1}$ і $S^2_{x_1x_2}$ або $S^2_{x_2}$ і $S^2_{x_1x_2}$, які можна об'єднати в зведену оцінку

$$\bar{s}'^2_{x_1x_2} = \frac{(u_1 - 1)s^2_{x_1} + (u_1 - 1)(u_2 - 1)s^2_{x_1x_2}}{(u_1 - 1) + (u_1 - 1)(u_2 - 1)} \approx \sigma^2_\varepsilon + m\sigma^2_{x_1x_2} \qquad (4.61)$$

або

$$\bar{s}''^2_{x_1x_2} = \frac{(u_2 - 1)s^2_{x_2} + (u_1 - 1)(u_2 - 1)s^2_{x_1x_2}}{(u_2 - 1) + (u_1 - 1)(u_2 - 1)} \approx \sigma^2_\varepsilon + m\sigma^2_{x_1x_2} \qquad (4.62)$$

з більшим числом ступенів свободи $\bar{\nu}'_{x_1x_2} = u_2(u_1 - 1)$, $\bar{\nu}''_{x_1x_2} = u_1(u_2 - 1)$.

В тому випадку, коли обидва дисперсійних відношення незначно відрізняються від одиниці, тобто вплив обох факторів несуттєвий ($\sigma^2_{x_1} = 0$ або $\sigma^2_{x_2} = 0$), однорідні оцінки $S^2_{x_1}$, $S^2_{x_2}$ і $S^2_{x_1x_2}$ для дисперсії $\sigma^2_\varepsilon + m\sigma^2_{x_1x_2}$ можна об'єднати в зведену оцінку

$$\bar{s}'''^2_{x_1x_2} = \frac{(u_1 - 1)s^2_{x_1} + (u_2 - 1)s^2_{x_2} + (u_1 - 1)(u_2 - 1)s^2_{x_1x_2}}{(u_1 - 1) + (u_2 - 1) + (u_1 - 1)(u_2 - 1)} \approx$$
$$\approx \sigma^2_\varepsilon + m\sigma^2_{x_1x_2} \qquad (4.63)$$

з більшим числом ступенів свободи $\bar{\nu}'''_{x_1x_2} = u_1u_2 - 1$.

2. Вплив взаємодії x_1x_2 з дисперсією

$$\sigma^2_{x_1x_2} \approx \frac{1}{m}\left(s^2_{x_1x_2} - \sigma^2_\varepsilon\right) \approx \frac{1}{m}\left(s^2_{x_1x_2} - s^2_\varepsilon\right) \qquad (4.64)$$

визнають суттєвим ($\sigma^2_{x_1x_2} > 0$), якщо відмінності $S^2_{x_1x_2}$ і S^2_ε будуть значними, тобто якщо

$$F_{12} = s^2_{x_1x_2} / s^2_\varepsilon > F_q\left(\nu_{x_1x_2}; \nu_\varepsilon\right). \qquad (4.65)$$

В іншому випадку вплив взаємодії вважається несуттєвим ($\sigma^2_{x_1x_2} = 0$) і тоді обидві однорідні оцінки $S^2_{x_1x_2}$ і S^2_ε для σ^2_ε можна об'єднати в одну зведену оцінку

$$\bar{s}_{\varepsilon}^{2}=\frac{(u_1-1)(u_2-1)s_{x_1x_2}^{2}+u_1u_2(m-1)s_{\varepsilon}^{2}}{(u_1-1)(u_2-1)+u_1u_2(m-1)}\approx\sigma_{\varepsilon}^{2} \qquad (4.66)$$

з більшим числом ступенів свободи $\bar{\nu}_{\varepsilon}=u_1u_2m-u_1-u_2+1$.

В тому випадку, коли вплив і факторів x_1, x_2, і їх взаємодія x_1x_2 несуттєві, дисперсію відтворюваності σ_{ε}^{2} можна оцінити вибірковою загальною дисперсією S_0^2.

4. Розрахункові формули для сум. Для практичних обчислень сум зручно користуватися їх перетвореними за формулою **(4.36)** виразами при таких позначеннях:

1) суми спостережень відгуку Y_j по рядках і Y_g' по стовпцях

$$Y_j=\sum_{g=1}^{u_2}\sum_{l=1}^{m}y_{jgl};\quad Y_g'=\sum_{j=1}^{u_1}\sum_{l=1}^{m}y_{jgl}; \qquad (4.67)$$

2) сума квадратів всіх спостережень

$$Q_1=\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\sum_{l=1}^{m}y_{jgl}^{2}; \qquad (4.68)$$

3) сума квадратів підсумків (сум) по рядках, розділена на число спостережень в рядку,

$$Q_2=\frac{1}{u_2m}\sum_{j=1}^{u_1}Y_j; \qquad (4.69)$$

4) сума квадратів підсумків (сум) по стовпцях, розділена на число спостережень в стовпці,

$$Q_3=\frac{1}{u_1m}\sum_{g=1}^{u_2}Y_g'^{2}; \qquad (4.70)$$

5) сума квадратів підсумків (сум) по серіях, розділена на число спостережень в серії,

$$Q_4=\frac{1}{m}\sum_{j=1}^{u_1}\sum_{g=1}^{u_2}\left(\sum_{l=1}^{m}y_{jgl}\right)^{2}; \qquad (4.71)$$

6) квадрат загального підсумка (суми), розділений на число всіх спостережень,

$$Q_5=\frac{1}{M}\left(\sum_{j=1}^{u_1}Y_j\right)^{2}=\frac{1}{M}\left(\sum_{g=1}^{u_2}Y_g'\right)^{2}. \qquad (4.72)$$

При цьому суми квадратів відхилень визначаються співвідношеннями

$$S_0=Q_1-Q_5;\quad S_{\varepsilon}=Q_1-Q_4;\quad S_{x_1}=Q_2-Q_5;$$

$$S_{x_2} = Q_3 - Q_5; \; S_{x_1x_2} = Q_4 + Q_5 - Q_2 - Q_3. \qquad (4.73)$$

Ми описали процедуру двофакторного дисперсійного аналізу. При багатофакторному аналізі послідовність операцій аналогічна, але значно ускладнюються таблиці спостережень і розрахункові формули.

ЗАВДАННЯ І ПОРЯДОК ВИКОНАННЯ РОБОТИ

1. Задати модель об'єкта №10 з одним фактором x_1 і середню квадратичну помилку відтворюваності експерименту σ_y=20.
2. Методом однофакторного ДА оцінити істотність впливу фактора x_1 на відгук y, для чого:
   - по черзі встановлювати довільні значення u=4 рівнів x_1 у діапазоні значень від 10 до 40;
   - на кожному j-у рівні x_1 провести серію з m=8 паралельних вимірювань величини відгуку y: y_{j1}, y_{i2}, ..., y_{jm};
   - отриманий експериментальний матеріал досліджувати за схемою однофакторного ДА.
3. Задати модель об'єкта №11 з двома факторами x_1 і x_2 і середню квадратичну помилку відтворюваності експерименту σ_y=20
4. Методом двофакторного ДА оцінити істотність впливу факторів x_1, x_2 і їхньої взаємодії x_1x_2 на відгук y, для чого:
   - по черзі встановлювати всі можливі комбінації довільних значень u_1= u_2= 4 рівнів x_1 і x_2 у діапазоні їх значень від 10 до 40;
   - для кожної комбінації рівнів x_1 і x_2 провести серію з m=4 рівнобіжних вимірювань величини відгуку y: y_{jg1}, y_{jg2}, ... , y_{jgm};
   - отриманий експериментальний матеріал досліджувати за схемою двофакторного ДА.

ФОРМУЛИ ДЛЯ РОЗРАХУНКУ

1. $\bar{y}_i = \dfrac{1}{m}\sum_{l=1}^{m} y_{il}$. 2. $\bar{\bar{y}} = \dfrac{1}{um}\sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl} = \dfrac{1}{u}\sum_{j=1}^{u}\bar{y}_j$.

3. $S_0 = \sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl}^2 - \dfrac{1}{um}\left(\sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl}\right)^2$. 4. $S_\varepsilon = \sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl}^2 - \dfrac{1}{m}\sum_{j=1}^{u}\left(\sum_{l=1}^{m} y_{jl}\right)^2$.

5. $S_x = \dfrac{1}{m}\sum_{j=1}^{u}\left(\sum_{l=1}^{m} y_{jl}\right)^2 - \dfrac{1}{um}\left(\sum_{j=1}^{u}\sum_{l=1}^{m} y_{jl}\right)^2$.

6. $S_0^2\{y\} = \dfrac{S_0}{um-1}$ с числом ступенів свободи $\upsilon_0 = um - 1$.

7. $S_{\varepsilon}^{2}\{y\} = \dfrac{S_{\varepsilon}}{u(m-1)}$ з числом ступенів свободи $v_{\varepsilon} = u(m-1)$.

8. $S_{x}^{2}\{y\} = \dfrac{S_{x}}{u-1}$ з числом ступенів свободи $v_{x} = u-1$.

9. $S_0 = S_{\varepsilon} + S_x$. 10. $v_0 = v_{\varepsilon} + v_x$. 11. $F = S_x^2 \big/ S_{\varepsilon}^2 > F_q(\nu_x; \nu_{\varepsilon})$.

12. $\bar{y}_{jg} = \dfrac{1}{m}\sum\limits_{l=1}^{m} y_{jgl}$. 13. $\bar{y}_j = \dfrac{1}{u_2 m}\sum\limits_{g=1}^{u_2}\sum\limits_{l=1}^{m} y_{jgl} = \dfrac{1}{u_2}\sum\limits_{g=1}^{u_2}\bar{y}_{jg}$.

14. $\bar{y}'_g = \dfrac{1}{u_1 m}\sum\limits_{j=1}^{u_1}\sum\limits_{l=1}^{m} y_{jgl} = \dfrac{1}{u_1}\sum\limits_{j=1}^{u_1}\bar{y}_{jg}$.

15. $\bar{\bar{y}} = \dfrac{1}{u_1 u_2 m}\sum\limits_{j=1}^{u_1}\sum\limits_{g=1}^{u_2}\sum\limits_{l=1}^{m} y_{jgl} = \dfrac{1}{u_1 u_2}\sum\limits_{j=1}^{u_1}\sum\limits_{g=1}^{u_2}\bar{y}_{jg} = \dfrac{1}{u_1}\sum\limits_{j=1}^{u_1}\bar{y}_j = \dfrac{1}{u_2}\sum\limits_{g=1}^{u_2}\bar{y}'_g$.

16. $Y_j = \sum\limits_{g=1}^{u_2}\sum\limits_{l=1}^{m} y_{jgl}$. 17. $Y'_g = \sum\limits_{j=1}^{u_1}\sum\limits_{l=1}^{m} y_{jgl}$. 18. $Q_1 = \sum\limits_{j=1}^{u_1}\sum\limits_{g=1}^{u_2}\sum\limits_{l=1}^{m} y_{jgl}^2$.

19. $Q_2 = \dfrac{1}{u_2 m}\sum\limits_{j=1}^{u_1} Y_j^2$. 20. $Q_3 = \dfrac{1}{u_1 m}\sum\limits_{g=1}^{u_2} Y'^2_g$. 21. $Q_4 = \dfrac{1}{m}\sum\limits_{j=1}^{u_1}\sum\limits_{g=1}^{u_2}\left(\sum\limits_{l=1}^{m} y_{jgl}\right)^2$.

22. $Q_5 = \dfrac{1}{u_1 u_2 m}\left(\sum\limits_{j=1}^{u_1} Y_j\right)^2 = \dfrac{1}{u_1 u_2 m}\left(\sum\limits_{g=1}^{u_2} Y'_g\right)^2$. 23. $S_0 = Q_1 - Q_5$.

24. $S_{\varepsilon} = Q_1 - Q_4$. 25. $S_{x_1} = Q_2 - Q_5$. 26. $S_{x_2} = Q_3 - Q_5$.

27. $S_{x_1 x_2} = Q_4 + Q_5 - Q_2 - Q_3$.

28. $S_0^2\{y\} = \dfrac{S_0}{u_1 u_2 m - 1}$ з числом ступенів свободи $v_0 = u_1 u_2 m - 1$.

29. $S_{\varepsilon}^2\{y\} = \dfrac{S_{\varepsilon}}{u_1 u_2 (m-1)}$ з числом ступенів свободи $v_{\varepsilon} = u_1 u_2 (m-1)$.

30. $S_{x_1}^2\{y\} = \dfrac{S_{x_1}}{u_1 - 1}$ з числом ступенів свободи $v_{x_1} = u_1 - 1$.

31. $S_{x_2}^2\{y\} = \dfrac{S_{x_2}}{u_2 - 1}$ з числом ступенів свободи $v_{x_2} = u_2 - 1$.

32. $S_{x_1 x_2}^2\{y\} = \dfrac{S_{x_1 x_2}}{(u_1 - 1)(u_2 - 1)}$ з числом ступенів свободи $v_{x_1 x_2} = (u_1 - 1)(u_2 - 1)$.

33. $S_0 = S_{\varepsilon} + S_{x_1} + S_{x_2} + S_{x_1 x_2}$. 34. $v_0 = v_{\varepsilon} + v_{x_1} + v_{x_2} + v_{x_1 x_2}$.

35. $F_1 = S^2_{x_1} \Big/ S^2_{x_1x_2} > F_q\left(\nu_{x_1};\nu_{x_1x_2}\right)$. 36. $F_2 = S^2_{x_2} \Big/ S^2_{x_1x_2} > F_q\left(\nu_{x_2};\nu_{x_1x_2}\right)$.

37. $F_{12} = S^2_{x_1x_2} \Big/ S^2_{\varepsilon} > F_q\left(\nu_{x_1x_2};\nu_{\varepsilon}\right)$.

КОНТРОЛЬНІ ЗАПИТАННЯ

1. Що називається факторами мінливості і випадковості?
2. Якого типу практичні задачі звичайно вирішують методом ДА?
3. Як математично формулюється задача однофакторного ДА?
4. Які основні передумови застосування ДА?
5. У чому полягає основна ідея методу ДА?
6. Яким чином проводиться кількісне оцінювання впливу факторів мінливості?
7. На які складові суми розкладається "загальна" сума квадратів відхилень у однофакторному ДА й вплив яких факторів вони характеризують?
8. На які складові суми розкладається "загальна" сума квадратів відхилень у двофакторному ДА й вплив яких факторів вони характеризують?
9. Спільною оцінкою яких генеральних дисперсій у однофакторному ДА є вибіркові дисперсії розсіювань: "загального", "у середині серій", "між серіями"?
10. Спільними оцінками яких генеральних дисперсій у двофакторному ДА є вибіркові дисперсії розсіювань: "загального", "у середині серій", "між рядками", "між стовпцями"?
11. Як у однофакторному ДА формуються вибіркові дисперсії розсіювань: "загального", "у середині серій", "між серіями"?
12. Як у двофакторному ДА формуються вибіркові дисперсії розсіювань: "загального", "у середині серій", "між рядками", "між стовпцями"; "між серіями"?
13. Яким чином проводиться оцінювання істотності впливу фактора мінливості у однофакторному ДА?
14. Яким чином проводиться оцінювання істотності впливу двох факторів мінливості і їхньої взаємодії у двофакторному ДА?

Додаток А

Таблиця А.1 – Таблиця рівномірно розподілених випадкових чисел у проміжку від 0 до 99

1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25*
10	09	73	25	33	76	52	01	35	86	34	67	35	48	76	80	95	90	91	17	39	29	27	49	49
37	54	20	48	05	64	89	47	42	96	24	80	52	40	37	20	63	61	04	02	00	82	29	16	65
08	42	26	89	53	19	64	50	93	03	23	20	90	25	60	15	95	33	47	64	35	08	03	36	06
99	01	90	25	29	09	37	67	07	15	38	31	13	11	65	88	67	67	43	97	04	43	62	76	59
12	80	79	99	70	80	15	73	61	47	69	03	23	66	53	98	95	11	68	77	12	17	14	68	33
66	06	57	47	17	34	07	27	68	50	36	69	73	61	70	65	81	33	98	85	11	19	92	91	70
31	06	01	08	05	45	57	18	24	06	35	30	34	26	14	86	79	90	74	39	23	40	30	97	32
85	26	97	76	02	02	05	16	56	92	68	66	57	48	18	73	05	38	52	47	18	62	38	85	79
63	57	33	21	35	05	32	54	70	48	90	55	35	75	48	28	46	82	87	09	83	49	12	56	24
73	79	64	57	53	03	52	96	47	78	35	80	83	42	82	60	93	52	03	44	35	27	38	84	35
98	52	01	77	67	14	90	56	86	07	22	10	94	05	58	60	97	09	34	33	50	50	07	37	98
11	80	50	54	31	39	80	82	77	32	50	72	56	82	48	29	40	52	42	01	52	77	56	78	51
83	45	20	96	34	06	28	89	80	83	18	74	67	00	78	18	47	54	06	10	68	71	17	78	17
88	68	54	02	00	86	50	75	84	01	36	76	66	79	51	90	36	47	64	93	29	60	91	10	62
99	59	46	73	48	37	51	76	49	69	91	82	60	89	28	93	78	56	13	68	23	47	83	41	13
65	48	11	76	74	17	46	85	09	50	58	04	77	69	74	73	03	95	71	86	40	21	81	65	44
80	12	43	56	35	17	72	70	70	15	45	31	82	23	74	21	11	57	82	53	14	38	55	37	63
74	35	09	98	17	77	45	27	72	14	43	23	60	02	10	45	52	16	42	37	96	28	60	26	55
69	91	62	68	03	66	25	22	91	48	36	93	68	72	03	76	62	11	39	90	94	40	05	64	18
09	89	32	05	05	14	22	56	85	14	46	42	75	67	88	96	29	77	88	22	54	38	21	45	98
91	49	91	45	23	68	47	92	76	86	46	16	28	35	54	94	75	08	99	23	37	08	98	00	48
80	33	69	45	98	26	94	03	68	58	70	29	73	41	35	53	14	03	33	40	42	05	08	23	41
44	10	48	19	49	85	15	74	79	54	32	97	92	65	75	57	60	04	08	21	22	22	20	64	13
12	55	07	37	42	11	10	00	20	40	12	86	07	46	97	96	64	48	94	39	28	70	72	58	15

Продовження таблиці А.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
09	18	82	00	97	32	82	53	95	27	04	22	08	63	04	83	38	98	73	74	64	27	85	80	44
90	04	58	54	97	51	58	15	06	54	94	93	88	19	97	91	87	07	61	60	68	47	66	46	59
73	18	95	02	07	47	67	72	62	69	52	29	06	46	64	27	12	46	70	18	41	36	18	27	60
75	76	87	64	90	20	97	18	17	49	90	42	91	22	72	95	37	50	58	71	93	82	34	31	78
08	35	86	99	10	73	64	24	27	85	13	66	15	88	73	04	61	68	75	53	31	22	30	84	20
28	30	60	32	64	81	33	31	05	91	40	51	00	78	93	32	60	46	04	75	94	11	90	18	40
53	84	08	62	33	81	59	41	36	28	51	21	59	02	90	28	46	66	87	95	77	76	22	07	97
91	75	75	77	41	61	61	36	22	69	50	26	39	02	12	65	78	17	65	14	83	48	34	70	65
89	41	59	26	94	00	39	75	83	91	12	60	71	76	46	48	94	97	23	06	94	54	13	74	08
77	51	30	38	22	86	83	42	89	01	58	41	48	27	74	51	90	81	39	80	72	89	35	57	07
19	50	23	71	74	69	97	92	02	88	55	21	02	97	73	34	28	77	52	51	65	34	46	75	15
21	81	85	99	13	93	27	88	17	67	05	68	67	31	56	07	08	28	50	46	31	85	33	84	52
51	47	46	64	99	68	10	72	36	21	94	04	99	13	45	42	83	60	91	91	08	00	74	64	49
99	65	96	83	31	62	53	52	41	70	69	77	71	28	30	74	81	97	81	42	43	86	07	28	34
33	71	34	80	07	93	68	47	28	69	51	92	66	47	21	58	30	32	98	22	93	17	49	39	72
85	27	49	68	93	11	30	32	92	70	28	83	43	41	37	73	51	59	04	00	71	14	84	36	43
84	13	38	96	40	33	03	55	21	66	73	85	27	00	91	51	22	26	05	61	66	32	71	84	23
56	73	21	62	34	17	39	59	61	31	10	12	39	16	22	85	49	65	75	66	81	60	41	88	80
65	13	85	68	06	87	64	88	52	61	34	31	36	58	51	45	87	52	10	68	85	64	44	72	77
38	00	10	21	76	81	71	91	17	11	71	60	29	29	37	74	21	96	40	49	65	58	44	96	98
37	40	29	63	97	01	31	47	65	86	56	27	11	00	86	47	32	46	26	05	40	03	03	74	38
97	12	54	03	48	87	08	33	14	17	21	81	53	92	60	75	23	79	20	47	15	50	12	95	78
21	82	64	11	84	47	14	33	40	72	64	53	88	59	02	59	13	90	64	41	03	85	65	45	59
73	13	54	27	42	95	71	90	90	35	85	79	47	42	96	08	79	98	81	57	64	69	11	92	02
07	63	87	79	29	03	06	11	80	73	96	20	74	41	56	23	82	19	95	38	04	71	76	69	94
60	52	88	74	41	07	95	41	98	14	59	17	52	06	95	05	53	35	21	99	61	21	20	44	55

Продовження таблиці А.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
63	60	64	93	29	16	50	53	44	84	40	21	95	25	63	43	65	17	70	82	07	20	73	17	90
61	19	69	04	46	26	45	74	77	74	61	92	43	37	29	65	39	45	95	93	42	58	26	05	27
15	47	44	52	66	95	27	07	99	53	59	36	78	38	48	82	39	61	01	18	33	21	15	94	66
91	55	72	85	73	67	89	75	43	87	54	62	24	44	31	91	19	04	25	92	92	92	74	59	13
42	48	11	62	13	97	34	40	87	21	16	86	84	87	67	03	07	11	20	59	25	70	14	66	70
23	52	37	83	18	73	20	88	98	37	68	93	69	14	16	26	25	22	96	63	05	52	28	25	62
04	49	35	24	94	75	24	63	38	24	45	86	25	10	25	61	96	27	93	35	68	73	71	24	72
00	54	99	76	54	64	05	18	81	69	96	11	96	38	96	54	69	28	23	91	23	28	72	95	29
35	96	31	53	07	26	89	30	93	64	33	35	93	54	52	77	97	45	00	24	90	10	33	93	33
59	80	80	83	91	45	42	72	68	42	83	60	94	97	00	13	02	12	46	92	78	56	52	01	06
46	05	88	52	36	01	39	09	22	86	77	28	14	40	77	93	91	08	36	47	70	61	74	29	41
32	17	90	05	97	89	37	92	52	41	05	56	70	70	07	86	74	31	71	57	95	39	41	18	38
69	23	46	14	06	20	11	74	52	04	15	95	66	00	00	18	74	39	24	23	97	11	89	63	38
19	56	54	14	30	01	75	87	53	79	40	41	52	15	85	66	67	43	68	06	84	96	28	52	07
45	15	51	49	38	19	47	60	72	46	43	66	79	45	43	59	04	79	00	33	20	82	66	95	41
94	86	43	19	94	36	16	81	08	51	34	88	88	15	53	01	54	03	54	56	05	51	45	11	76
98	08	62	48	26	45	24	02	84	04	44	99	90	88	96	39	09	47	34	07	35	44	13	18	80
33	18	51	62	32	41	94	15	09	49	89	43	54	85	81	88	69	54	19	94	37	54	87	30	43
80	95	10	04	06	96	38	27	07	74	20	15	12	33	87	25	01	52	52	98	94	62	46	11	71
79	75	24	91	40	71	96	12	82	96	69	86	10	25	91	74	85	22	05	39	00	38	75	95	79
18	63	33	25	37	98	14	50	65	71	31	01	01	46	74	05	45	56	14	27	77	83	89	19	39
74	02	94	39	02	77	55	73	22	70	97	79	01	71	19	52	62	75	80	21	80	81	45	17	48
54	17	84	56	11	80	99	33	71	43	05	33	51	29	69	56	12	71	92	55	36	04	09	53	24
11	66	44	98	83	52	07	98	48	27	59	38	17	15	39	09	97	33	34	40	88	46	12	33	56
48	32	47	79	28	31	24	96	47	10	42	29	53	68	70	32	30	75	75	46	15	02	00	99	94
69	07	49	41	38	87	63	79	19	76	35	58	40	44	01	10	51	82	16	15	01	84	87	69	38

Продовження таблиці А.1

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
83	59	53	56	55	06	95	89	29	83	05	12	80	97	19	77	43	35	37	83	92	30	15	04	98
10	85	06	26	46	99	59	91	05	47	13	49	90	63	19	53	07	57	18	39	06	41	01	93	62
39	82	09	89	52	43	62	26	31	47	64	42	18	08	14	43	80	00	93	51	31	02	47	31	67
59	58	00	64	78	75	56	97	88	00	88	83	55	44	86	23	76	80	61	56	04	11	10	84	08
38	50	80	73	41	28	79	34	87	63	90	82	29	70	22	17	71	90	42	07	95	95	44	99	53
30	69	27	06	68	94	68	81	61	27	56	19	68	00	91	82	06	76	34	00	05	46	26	92	00
65	44	39	56	59	18	28	82	74	37	49	63	22	40	41	08	33	76	56	76	96	29	99	08	36
27	26	75	02	64	13	19	27	22	94	07	47	74	46	06	17	98	54	89	11	97	34	13	03	58
91	30	70	69	91	19	07	22	42	10	36	69	95	37	28	28	82	53	57	93	28	97	66	62	52
68	43	49	46	88	84	47	31	36	22	62	12	69	84	08	12	84	38	25	90	09	81	59	31	46
48	90	81	58	77	54	74	52	45	91	35	70	00	47	54	83	82	45	26	92	54	13	05	51	60
06	91	34	51	97	42	67	27	86	01	11	88	30	95	28	63	01	19	89	01	14	97	44	03	44
10	45	51	60	19	14	21	03	37	12	91	34	23	78	21	88	32	68	08	51	43	66	77	08	83
12	88	39	73	43	65	02	76	11	84	04	28	50	13	92	17	97	41	50	77	90	71	22	67	69
21	77	83	09	76	38	80	73	69	61	31	64	94	20	96	63	28	10	20	23	08	81	64	74	49
19	52	85	95	15	65	12	25	96	59	86	28	36	82	58	69	59	21	37	98	16	43	59	15	29
67	24	55	26	70	35	58	31	65	63	79	24	68	66	86	76	46	33	42	22	26	65	59	08	02
60	58	44	73	77	67	50	03	79	92	45	13	42	65	29	26	76	68	36	37	41	32	64	43	44
53	85	34	13	17	36	06	69	48	50	58	83	87	38	59	49	36	47	33	31	96	24	04	36	42
24	63	73	87	36	74	38	48	93	42	52	62	30	79	92	12	36	91	86	01	03	74	28	38	73
83	08	01	24	51	38	99	22	28	25	07	75	95	17	77	97	37	72	75	85	51	97	23	78	67
16	44	42	43	34	36	15	19	20	73	27	49	37	09	39	85	13	03	25	52	54	84	65	47	59

Додаток Б

Таблиця Б. 1 – Верхні односторонні границі для величини $G_{кр}$ в залежності від числа ступенів свободи чисельника (ν_1) та знаменника (ν_2) для G-розподілення Кохрена при рівні значущості q = 0,05

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	16	36	144	∞
2	0,9985	0,9750	0,9392	0,9057	0,8584	0,8534	0,8332	0,8159	0,8010	0,7880	0,7341	0,6602	0,5813	0,5000
3	9669	8709	7977	7457	7071	6771	6530	6333	6167	6025	5466	4748	4031	3333
4	9065	7679	6841	6287	5895	5598	5365	5175	5017	4884	4366	3720	3093	2500
5	0,8412	0,6838	0,5981	0,5440	0,5063	0,4783	0,4564	0,4387	0,4241	0,4118	0,3645	0,3066	0,2513	0,2000
6	7808	6161	6321	4803	4447	4148	3980	3817	3682	3568	3135	2612	2119	1667
7	7271	5612	4800	4307	3907	3726	3555	3384	3254	3154	2756	2273	1833	1429
8	0,6798	0,5157	0,4377	03910	03595	0,3362	0,3185	0,3043	0,2926	0,2829	0,2462	0,2020	0,1516	0,1250
9	6385	4775	4027	3584	3586	3067	2901	2768	2659	2568	2226	1820	1446	1111
10	6020	4450	3733	3311	3029	2823	2666	2541	2439	2353	2032	1655	1308	1000
12	0,6410	0,3924	0,3264	02880	0,2624	0,2439	0,2299	0,2187	0,2098	0,2020	0,1737	0,1403	0,1100	0,0833
15	4709	3346	2758	2419	2195	2034	1911	1315	1736	1671	1429	1144	0889	0667
20	3894	2705	2205	1921	1835	1602	1601	1422	1357	1303	1108	0879	0675	0500
24	0,3434	0,2354	0,1907	0,1656	0,1493	0,1374	0,1286	0,1216	0,1160	0,1113	0,0942	0,0743	0,0567	0,0417
30	2929	1980	1593	1377	1237	1137	1061	1002	0958	0921	0771	0604	0457	0337
40	2370	1576	1259	1082	0968	0887	0827	0780	0745	0713	0595	0462	0347	0250
60	0,1737	0,1131	0,0895	0,0766	0,0682	0,0623	0,0583	0,0552	0,0520	0,0487	0,0411	0,0316	0,0234	0,0167
120	0998	0632	0495	0419	0371	0337	0312	0292	0279	0266	0218	0165	0120	0083
∞	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000	0000

Додаток В

Таблиця В.1 – Двосторонні границі для величини $t_{кр}$ в залежності від числа ν ступенів свободи та від рівня значущості q-імовірності $P\{t > t_{кр}\}$ для t-розподілення Стьюдента

ν \ q	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
1	3,077	6,313	12,706	31,820	63,656	127,656	318,308	636,619
2	1,885	2,920	4,302	6,964	9,924	14,089	22,327	31,599
3	6377	3534	3,182	4,540	5,840	7,458	10,214	12,924
4	5332	1318	2,776	3,746	4,604	5,597	7,173	8,610
5	4759	0150	5706	3649	0321	4,773	5,893	6,869
6	1,439	1,943	2,446	3,142	3,707	4,316	5,207	5,958
7	4149	8946	3646	2,998	4995	0293	4,785	4079
8	3968	8595	3060	8965	3554	3,832	5008	0413
9	3830	8331	2622	5214	2498	6897	2968	4,780
10	3720	8125	2281	7638	1693	5814	1437	5869
11	1,363	1,796	2,201	2,718	3,105	3,496	4,024	4,437
12	3562	7823	1788	6810	0545	4284	3,929	3178
13	3502	7709	1604	6503	0123	3725	8520	2208
14	3450	7613	1448	6245	2,976	3257	7874	1405
15	3406	7530	1314	6025	9467	2860	7328	0728
16	1,336	1,746	2,119	2,583	2,920	3,252	3,686	4,015
17	3334	7396	1098	5668	8982	2224	6458	3,965
18	3304	7341	1009	5514	8784	1966	6105	9216
19	3277	7291	0930	5395	8609	1737	5794	8834
20	3253	7247	0860	5280	8453	1534	5518	8495
21	1,323	1,721	2,079	2,517	2,831	3,135	3,527	3,819
22	3212	7167	0739	5083	8188	1188	5050	7921
23	3195	7139	0687	4999	8073	1040	4850	7676
24	3178	7109	0639	4922	7969	0905	4668	7454

Продовження таблиці В.1

ν \ q	0,20	0,10	0,05	0,02	0,01	0,005	0,002	0,001
25	3163	7081	0595	4851	7874	0782	4502	7251
26	1,315	1,706	2,055	2,478	2,778	3,066	3,435	3,706
27	3137	7033	0518	4727	7707	0565	4210	6896
28	3125	7011	0484	4671	7633	0469	4082	6739
29	3114	6991	0452	4620	7564	0380	3962	6594
30	3104	6973	0423	4573	7500	0298	3852	6466
32	1,308	1,693	2,036	2,448	2,738	3,014	3,365	3,621
34	3070	9609	0322	4411	7284	0020	3479	6007
36	3050	6883	0281	4345	7195	2,990	3326	5821
38	3042	6860	0244	4286	7116	9808	3190	5657
40	3031	6839	0211	4233	7045	9712	3069	5510
42	1,320	1,682	2,018	2,418	2,698	2,693	3,296	3,537
44	3011	6802	0154	4141	6923	9555	2861	5258
46	3002	6787	0129	4102	6870	9488	2771	5150
48	2994	6772	0106	4066	6822	9426	2669	5051
50	2987	6759	0086	4033	6778	9370	2614	4960
55	1,997	1,673	2,004	2,396	2,668	2,924	3,256	3,476
60	2958	6706	0003	3901	6603	9146	2317	4602
65	2947	6686	1,997	3851	6536	6060	2204	4466

Додаток Г

Таблиця Г.1 – Верхні односторонні границі $F_{кр}$ в залежності від чисел ступенів свободи чисельника (ν_1) та знаменника (ν_2) для F-розподілення Фішера при рівні значущості q=0,05

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9
1	161	200	216	225	230	234	237	239	241
2	18,51	19,00	19,16	19,25	19,30	19,33	19,36	19,37	19,38
3	10,13	9,55	9,28	9,12	9,01	8,94	8,88	8,84	8,81
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,78
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10
7	5,59	4,74	4,35	4,12	3,97	3,87	3,679	3,73	3,68
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90
12	4,75	3,88	3,49	3,26	3,11	300	2,92	2,85	2,80
13	4,67	3,80	3,41	3,18	3,02	2,92	2,84	2,77	2,72
14	4,60	3,74	3,34	3,11	2,96	2,85	2,77	2,70	2,65
15	4,54	3,68	3,29	3,06	2,90	2,79	2,70	2,64	2,59
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54
17	4,45	3,59	3,20	2,96	2,81	2,70	2,62	2,55	2,50
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46
19	4,38	3,52	3,13	2,90	2,74	2,63	2,55	2,48	2,43
20	4,35	3,49	3,10	2,87	2,71	2,60	2,52	2,45	2,40
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37
22	4,30	3,44	3,05	2,82	2,66	2,55	2,47	2,40	2,35
23	4,28	3,42	3,03	2,80	2,64	2,53	2,45	2,38	2,32
24	4,26	3,40	3,01	2,78	2,62	2,51	2,43	2,36	2,30
25	4,24	3,38	2,99	2,76	2,60	2,49	2,41	2,34	2,28
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,30	2,25
28	4,20	3,34	2,95	2,71	2,56	2,44	2,36	2,29	2,24
29	4,18	3,33	2,93	2,70	2,54	2,43	2,35	2,28	2,22
30	4,17	3,32	2,92	2,69	2,53	2,42	2,34	2,27	2,21
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12
60	4,00	3,15	2,76	2,52	2,37	2,25	2,17	2,10	2,04
120	3,92	3,07	2,68	2,44	2,29	2,17	2,08	2,01	1,95
∞	3,84	2,99	2,60	2,37	2,21	2,09	2,01	1,94	1,88

Продовження таблиці Г.1

ν_2 \ ν_1	10	12	16	20	24	30	40	60	120	∞
1	242	244	246	248	249	250	251	253	253	254
2	19,36	19,41	19,43	19,44	19,45	19,46	19,47	19,48	19,49	19,50
3	8,78	8,74	8,69	8,66	8,64	8,62	8,60	8,57	8,56	8,53
4	5,96	5,91	5,84	5,80	5,77	5,74	5,71	5,68	5,66	5,63
5	4,47	4,68	4,60	4,56	4,53	4,50	4,46	4,42	4,40	4,36
6	4,06	4,00	3,92	3,87	3,84	3,81	3,77	3,72	3,71	3,67
7	3,63	3,57	3,49	3,44	3,41	3,38	3,34	3,279	3,28	3,23
8	3,34	3,28	3,20	3,15	3,12	3,08	3,05	3,00	2,98	2,93
9	3,13	3,07	2,98	2,93	2,90	2,86	2,82	2,77	2,76	2,71
10	2,97	2,91	2,82	2,77	2,74	2,70	2,67	2,61	2,59	2,54
11	2,86	2,79	2,70	2,65	2,61	2,57	2,53	2,47	2,45	2,40
12	2,76	2,69	2,60	2,54	2,50	2,46	2,42	2,36	2,35	2,30
13	2,67	2,60	2,51	2,46	2,42	2,38	2,34	2,28	2,26	2,21
14	2,60	2,53	2,44	2,39	2,35	2,31	2,27	2,21	2,19	2,13
15	2,55	2,48	2,39	2,33	2,29	2,25	2,21	2,15	2,12	2,07
16	2,49	2,42	2,33	2,28	2,24	2,20	2,16	2,09	2,07	2,01
17	2,45	2,38	2,29	2,23	2,19	2,15	2,11	2,04	2,02	1,96
18	2,41	2,34	2,25	2,19	2,15	2,11	2,07	2,00	1,98	1,92
19	2,38	2,31	2,21	2,15	2,11	2,07	2,02	1,96	1,94	1,88
20	2,35	2,28	2,18	2,12	2,08	2,04	1,89	1,92	1,90	1,84
21	2,32	2,25	2,15	2,09	2,05	2,00	1,96	1,89	1,87	1,81
22	2,20	2,23	2,13	2,07	2,03	1,98	1,93	1,87	1,84	1,78
23	2,28	2,20	2,10	2,04	2,00	1,96	1,91	1,84	1,82	1,76
24	2,26	2,18	2,09	2,02	1,98	1,94	1,89	1,82	1,80	1,73
25	2,24	2,16	2,06	2,00	1,96	1,92	1,87	1,80	1,77	1,71
26	2,22	2,15	2,05	1,99	1,90	1,95	1,85	1,78	1,76	1,69
27	2,20	2,13	2,03	1,97	1,93	1,88	1,84	1,76	1,74	1,67
28	2,19	2,12	2,02	1,96	1,91	1,87	1,81	1,75	1,72	1,65
29	2,18	2,10	2,00	1,94	1,90	1,85	1,80	1,73	1,71	1,64
30	2,16	2,09	1,99	1,93	1,89	1,84	1,79	1,72	1,69	1,62
40	2,07	2,00	1,90	1,84	1,79	1,74	1,69	1,61	1,59	1,51
60	1,99	1,92	1,81	1,75	1,70	1,65	1,59	1,50	1,48	1,39
120	1,90	1,83	1,72	1,65	1,60	1,65	1,55	1,49	1,36	1,25
∞	1,83	1,75	1,64	1,57	1,52	1,46	1,39	1,28	1,24	1,00

Література

1. *Адлер Ю.П., Маркова Е.В., Грановский Ю.В.* Планирование эксперимента при поиске оптимальных условий. –М.: Наука, 1976.
2. *Иванов А.З., Круг Г.К., Филаретов Г.Ф.* Статистические методы в инженерных исследованиях. Регрессионный анализ. –М.: МЭИ, 1977.
3. *Иванов А.З., Круг Г.К., Филаретов Г.Ф.* Статистические методы в инженерных исследованиях. Планирование первого порядка для регрессионных экспериментов. –М.: МЭИ, 1978.
4. *Налимов В.В., Чернова Н.А.* Статистические методы планирования экстремальных экспериментов. –М.: Наука, 1965.
5. *Растригин Л.А.* Статистические методы поиска. –М.: Наука, 1968.
6. *Мойсюк Б.Н.* Элементы теории оптимального эксперимента. Ч.II. –М.: МЭИ, 1976.
7. *Ермаков С.М., Жиглявский А.А.* Математическая теория оптимального эксперимента. –М.: Наука, 1987.
8. Статистические методы в инженерных исследованиях./ Под. ред. *Круга Г.К.* –М.: Высшая школа, 1983.
9. Исследование устройств автоматики методом планирования эксперимента./ Под ред. *Воронова В.Г.* –Харьков: Вища школа, 1986.
10. *Фролов В.Н., Львович Я.Е., Меткин Н.П.* Автоматизированное проектирование технологических процессов и систем производства РЭС. –М.: Высшая школа, 1991.

Навчальне видання

Мотигін Володимир В'ячеславович
Павлов Сергій Миколайович

ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ В ІНЖЕНЕРНИХ ДОСЛІДЖЕННЯХ

(ЛАБОРАТОРНИЙ ПРАКТИКУМ)

Навчальний посібник

Оригінал-макет підготовлено авторами

Редактор С.А.Малішевська

Підписано до друку
Формат 29,7х42 $1/4$ Гарнітура Times New Roman
Друк різографічний Ум. друк. арк. 4,83
Тираж 75 прим.
Зам. № 2000-0127

Віддруковано в комп'ютерному інформаційно-видавничому центрі Вінницького державного технічного університету
21021, м. Вінниця, Хмельницьке шосе, 95, ВДТУ, ГНК, 9-й поверх
Тел. **(0432) 44-01-59**